RICE UNIVERSITY

Semantic and Syntactic Interference in Sentence Comprehension and Their Relationship to Working Memory Capacity

by

Yingying Tan

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Arts

APPROVED, THESIS COMMITTEE:

Randi Martin, Professor,
Psychology

David Lane, Professor,
Psychology

Tatiana Schnur, Assistant Professor,
Psychology

HOUSTON, TEXAS
DEC 2011

ABSTRACT

Semantic and Syntactic Interference in Sentence Comprehension and Their Relationship to Working Memory Capacity

by

Yingying Tan

This study investigated the nature of the relationship between working memory (WM) and sentence processing by examining interference effects in sentence comprehension and relating those to performance on a set of WM tasks, executive function tasks, and vocabulary tests. For online sentence comprehension, semantic interference effects were negatively correlated with semantic retention capacity. Syntactic interference effects were negatively related only to reading span. These results are consistent with the multiple capacities account (Martin & Romani, 1994), which postulates that there are separable retention abilities for semantic, syntactic, and phonological information, with the first two being critical for sentence comprehension. For offline sentence comprehension, participants with better semantic STM, WM span, vocabulary, or Stroop performance showed less difficulty in semantic interference resolution. These results were consistent to some extent with multiple capacities account, the general resources account (Just & Carpenter, 1992) and retrieval-based interference account (Van Dyke 2007).

Keywords: Interference effect, Working memory capacity, Cue-based retrieval, Sentence processing

Acknowledgements

First I would like to show my sincerest gratitude and appreciation to my supervisor, Dr. Randi Martin, who has supported me through my thesis with her patience, understanding, and vast knowledge. This work would not have been possible without the guidance from her.

I would like to thank Dr. Julie Van Dyke, for generously sharing her experimental materials and experience with me. I am also grateful to the help and guidance from all the other committee members: Dr. David Lane, for providing me with statistical advice, and Dr. Tatiana Schnur, for her insightful comments and questions.

I owe my special thanks to my parents and my boyfriend Yiyang Shen for continuous and unconditional support. Their supporting helped me through the difficult times.

Last but not the least, thanks to all my friendly and thoughtful lab members, Loan Voung, Corinne Allen, Azli Hassen, Yi Guo, for their generous help and providing a cheerful lab environment. Also I thank all the research assistants who helped me with my experiment: Sarah Zelonis, Sanam Jivani, Maria Pickett, and Daniel Machao.

Table of Contents

List of Figures and Tables

1 Introduction

1.1 Working Memory and Sentence Comprehension

Understanding spoken or written language in real time is essential to our daily life. A number of studies have shown a link between working memory processes and language comprehension. The general implication drawn from these findings is that working memory supports language processing. However, researchers have not come to a final conclusion about the nature of this relationship, such as exactly how working memory is involved in language processing, or whether there is a language-specific working memory system. On one account, if WM capacity limits language comprehension, then people with smaller WM capacity should experience greater constraints (Daneman & Carpenter, 1980; Just & Carpenter, 1992; Gordon, Hendrick & Johnson, 2001; Gordon, Hendrick, Johnson & Lee, 2006). Considerable evidence shows that WM capacity tapped by complex span measures (e.g., reading span, operation span, etc.) is a good predictor of people's performance in sentence comprehension (Daneman & Carpenter, 1980; Fedorenko, Gibson & Rohde, 2006, 2007; Just & Carpenter, 1992; Gordon et al., 2001, 2002; Kane et al., 2001). On another account, researchers have implied that capacity is not critical; instead, the ability to retrieve relevant information may be critical (Lewis, 1996, Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006; Van Dyke, 2007). Such researchers have adopted current views of WM which claim that the number of items that can be maintained in the focus of attention is extremely limited for all types of tasks (McElree, 2006), including sentence processing (Lewis, 1996). That is, no more than two items may be in the focus of attention, which is within everyone's capacity. According to this view, individual differences could arise in differences in the ability to retrieve information outside of the focus - either due to variation across individuals

in the quality of the encoded information, the ability to develop appropriate retrieval cues, or the ability to resist interference from information that partially matches retrieval cues (Van Dyke et al., 2010).

Based on the previous discussion, the questions we try to answer in the current project is that: a) What is the nature of the capacity limit during sentence comprehension – that is, is it due to storage capacity or differences in cue-based retrieval? b) If sentence processing is constrained by WM capacity, does sentence comprehension recruit the same WM resources used by more general cognitive processes, or is there a specific WM resource underlie sentence comprehension? We first review the evidence for limited WM capacity account, and then consider the interference account and its supporting evidence.

1.2 Storage-based Account

According to the *capacity* account, there should be a positive correlation between WM capacity and people's performance on language processing. Early research on WM focused on the ability to store items for later retrieval after a brief interval and used simple span tasks like digit span as the measure. Early studies from Baddeley (1984) and many others (e.g., Conrad, 1964) found that traditional verbal memory span tasks rely on phonological short-term storage. Such storage might be involved during both oral and written sentence comprehension (Caramazza et al., 1981; Saffran & Marin, 1976). However, Just and Carpenter (1992) suggested that the role working memory capacity plays in sentence comprehension is to provide storage for the partial products of comprehension, like multilevel representation of earlier words and phrases, theme of the text and so on, and to provide support for the processing the rest of the sentence. They suggested that individual differences in working memory capacity should be expressed as the maximum amount of activation

available to support the combination of storage or processing. Thus they argued that people varied in the number of features that they could maintain and manipulate. Complex span tasks have been developed to tap both the storage capacity and the processing component of WM (Daneman & Carpenter, 1980; Engle, Tuholski, et al., 1999). Daneman and Carpenter (1980) developed a reading span measure that they argued reflects both storage and processing. In this task, subjects are required to read aloud or listen to a set of sentences, and recall the final word of each sentence at the end of each set. The set size typically varies from two to six sentences. Overall WM capacity is determined by totaling the number of words subjects recalled from all the correct trials. Their results showed that the number of final words people could recall in reading span correlated highly with other reading comprehension measures, such as verbal SAT.

There is controversy, however, regarding whether the WM capacity underlying language comprehension is domain-specific for sentence processing, or whether there is a domain-general capacity common to all verbal tasks.

1.2.1 Domain general resources account

One possibility advanced by Just and Carpenter and some other researchers is that there are general verbal processing resources that underlie all verbal tasks. The main evidence is based on the interactions between sentence processing, external memory load and working memory capacity (Gorden, Hendrick, & Johnson, 2001, 2004; Gorden, Hendrick, & Levine, 2002; Just & Carpenter, 1992; King & Just, 1991). In an early study, King and Just (1991) separated their subjects into a high-WM group and low-WM group according to their performance on the reading span measure. They found that subjects with lower spans showed longer reading times and lower accuracy on more demanding object-relative sentences (e.g., "The woman that the

man kissed had red hair") than high span subjects whereas the two groups performed similarly on less demanding subject-relative sentences (e.g., "The woman that kissed the man had red hair"). They suggested that working memory constraints are manifested when processing demands exceed capacity.

In addition, when the subjects were required to complete a dual task during sentence reading, their performance on a subsequent comprehension question declined compared to the condition without dual tasks, especially for the low span readers who had less capacity for the complex computations. In some recent studies, researchers manipulated syntactic complexity and difficulty of the dual tasks at the same time (Fedorenko, Gibson & Rohde, 2006; Fedorenko, Gibson, & Rohde, 2007; Gordon, Hendrick & Levine, 2002). In these experiment, subjects were required to read some sentences of varying syntactic complexity (e.g. containing object– or subject –extracted relative), while remembering a short set of words (Fedorenko et al., 2006; Gordon et al., 2002) or completing arithmetic integration or spatial integration based on memory (Fedorenko et al., 2007). The words to remember were either from the same categories as the sentence nouns or not. (common nouns vs. proper names). The results showed that the match between the words' category in the memory load and in the sentence impaired sentence comprehension on both comprehension question accuracy (Gordon et al., 2002) and self-paced reading time (Fedorenko et al., 2006). Additionally, the interactions between sentence processing and dual tasks were only observed when the secondary task is verbal based (e.g. words list or arithmetic operation), but not spatial based (Fedorenko et al., 2007). Therefore, the researchers concluded that there was a trade-off between storage and processing: the more working memory demands or the lower the capacity, the fewer the resources that are available for sentence processing. Accordingly, there will be more difficulties in

sentence comprehension. Such trade-off further confirmed that linguistic processing shared a pool of working memory recourses as other verbal working memory tasks.

Some researchers have argued against Just and Carpenter's domain-general or single-resource account by suggesting a separate sentence-interpretation resource (Caplan & Waters, 1996; Evans, Waters, & Caplan, 2010) or implying that only a certain component of WM is related to sentence processing ability (Martin & Romani, 1994; Martin & He, 2004). Evidence contrary to Just and Carpenter's claims was obtained from studies with healthy individuals which failed to show an interaction between WM span and sentence comprehension difficulty (Caplan & Waters, 1999) and studies of brain damaged patients in which patients with working memory deficits (very restricted STM span) showed excellent sentence comprehension (Butterworth et al., 1986; Caplan & Waters, 1999; Martin, 1987).

1.2.2 Domain specific resources account

Caplan and Waters (1999) suggested that there should be no relation between sentence comprehension and WM capacity because comprehension implicates dedicated capacities not tapped by regular span measures. Thus, general working memory capacities should not predict language processing efficiency. Caplan and Waters proposed a sentence-processing model with two processing procedures: interpretive and post-interpretive. The "interpretive processing" component consists of on-line processing including extracting meaning from a linguistic signal, like "recognizing words and appreciating their meanings and syntactic features, constructing syntactic and prosodic representations and assigning thematic roles, focus, and other aspects of propositional and discourse-level semantics". By "post-interpretive processing", they referred to the offline processes as "storing information in long-term semantic memory, reasoning, planning actions, and other functions". At

the sentence level, subjects assign the syntactic structure of a sentence and use that structure to determine the meaning of a sentence during interpretive processing, then make use of that meaning to accomplish other tasks (e.g. answering comprehension questions, matching sentence and pictures) in the post-interpretive processing. Based on this model, Caplan and Waters claimed that verbal working memory could be further divided into two sub-pools: verbal WM for interpretive processing and verbal WM for post-interpretive processing. Standard span tasks only tap post-interpretive processing, which involves using the products of interpretive processing to complete some task. This two-stage sentence processing model had important implications for sentence processing studies. Both researchers who support the domain-general resource account or the sentence-specific resource account agreed that the WM capacity tapped by span tasks is involved in offline sentence processing (e.g. answering a comprehension question). But it remained controversial whether a general WM capacity supports online sentence processing (e.g. self-paced reading or eye-tracking).

In support of their model, Caplan and Waters asserted that the poorer performance of people with low WM capacity is not necessarily due to inability to accomplish the processing but could be due to their difficulties in satisfying other aspects of task demands, like dividing their attention in a dual-task-experiment. They reanalyzed or replicated several previous studies and failed to find differences between different capacity groups in syntactic processing, to find effects of verbal memory load on syntactic processing, or to find difference between high-WM-span subjects and low-WM-span subjects of the effect of concurrent load on syntactic processing. They found that digit load only interfered with sentence complexity when syntactic processing was interrupted by a secondary task, such as when the

presentation of sentence was interrupted by a series of words while it is being presented or was staggered across the presentation of secondary task (as in sentence final word recall task), but not when the digit load was presented prior to the sentence and digit recall occurred after the sentence had been understood (Waters et al, 1987; Waters & Caplan, 1996). Thus, Caplan and Waters concluded that an interaction between load and sentence complexity is not caused by a shared resource pool for storing a sequence of digits and constructing a syntactic interpretation of a sentence, but because of the attentional shift between the two tasks when the load was presented concurrently. Span tasks share the same sources with operations on the propositional content of a sentence, but not with on-line syntactic processing.

Additionally, in response to Fedorenko and colleagues' recent study (2007) as I reviewed above, Evans, Waters, and Caplan (2010) replicated Fedorenko et al.'s experiments but failed to fully reproduce their results. Evan and colleagues did not find an interaction of sentence complexity and mathematical complexity in self-paced reading time, but only in total reading time of eye-tracking. However, Evans et al. suggested that such results did not solve the critical question of what processes are involved in the shared resource postulated by Fedorenko and colleagues. The interaction effect in eye tracking data occurred at the NP in the relative clause, rather than the verb. Previous studies have only attributed the effects of object-relative-structure on the verb in relative clause to integration of information retrieved from memory with incoming information. It is not so clear whether the significant interaction observed on NP was caused by shared WM resources. Neither the location nor the timing of the observed interaction suggested that it arises because of shared resources between syntactic and arithmetical operations. Thus, although Evans et al. agreed that such results require further study, they suggested that these results do

weaken the claims for a shared working memory system.

1.2.3 Multiple capacities account

Although Caplan and Waters put forward a convincing case about the separation of the WM resources involved in syntactic processing from those tapped by various span tasks such as reading span or digit span, their arguments regarding interpretive and post-interpretive processes and the relation to span tasks are less well justified. It is not clear how to use their theory to explain the effect on comprehension of number of propositions and its interaction with extraneous load since they included "assigning thematic roles, focus, and other aspects of propositional and discourse-level semantics" as part of the interpretive process (Martin, 1999). A different approach which relates some aspects of online sentence processing to the capacities tapped by span tasks was put forward by Martin and colleagues based on neuropsychological results (Hanten & Martin, 2000; Martin & Romani, 1994; Martin & He, 2004; Martin, 2003). In this multiple capacities approach, Martin et al. argued that verbal working memory was broken down into two separate capacities for the retention of phonological and semantic information. The phonological retention capacity is relevant to verbatim repetition and the learning new words, but is irrelevant to sentence comprehension. The semantic retention capacity is critical for language comprehension and production, as well as for learning new meanings.

Martin and colleagues obtained behavioral and neuronanatomical evidence for the double dissociation of these two retention capacities. Martin et al. (2003b) collected neuroimaging data from undergraduate students while they were doing a probe recognition task. Semantic retention ability was measured by manipulating memory load in a synonym judgment task (judge whether a probe word is a synonym of any list word), and phonological retention ability was measured by manipulating

memory load in a rhyme probe task (judge whether a probe word rhymed with any list item). The phonological STM manipulation caused more activation in the left inferior parietal region, while the semantic STM manipulation caused greater activation in a left frontal region.

Studies of brain-damaged patient in Martin's lab also supported the dissociation of semantic short-term memory and phonological short-term memory (Martin, Shelton & Yaffee, 1994; Martin & Romani, 1994; Martin & He, 2004). In an early study, Martin and Romani (1994) tested three brain-damaged patients. They compared patient (A.B.) who had a specific deficit in semantic retention with a patient (E.A.) who had a specific deficit in phonological retention and a second (M.W.) with a nearly normal STM capacity. To examine their sentence comprehension ability, the patients were tested on sentence anomaly judgments (judge whether a sentence make sense or not) and grammaticality judgments (judge whether a sentence is grammatically acceptable, e.g. *The girl accidentally cut himself while playing.*). In the anomaly judgment condition, the number of the nouns or the adjectives varied from one to three items, and the position of the nouns/adjectives was varied as before or after the corresponding verbs/nouns (e.g. before condition: "The rusty, old, red swimsuit was "... vs. after condition: "The swimsuit that was old, red, and rusty... "). Martin et al. proposed that the manipulation of words position affected the demand for the retention of individual word meanings because in the condition in which more than one nouns/adjectives preceded the verb/noun integration of the adjectives with the nouns or the nouns with the verb was delayed. In contrast, when the noun/adjectives appeared after the verb/noun, integration could occur immediately as each word was processed. In the memory-stressed grammaticality judgments, besides varying the grammatical structure of the sentences, Martin et al. also manipulated the

amount of intervening material between the words that made a structure ungrammatical to check subjects' ability to retain part of a grammatical structure for later integration. In the intervening condition words were added between the words signaling the ungrammaticality of the sentences and in the across-the-board condition words were added at other points of the sentences. The results showed that only patient A.B. (with a semantic STM deficit) showed a deficit in the anomaly judgment task that required holding the meaning of several individual words before integration, while the other two patients performed within normal range. However, in the grammaticality judgment test, only patient M.W. showed a significant overall difference between the intervening and across-the-board conditions whereas A.B. did not. This pattern is the reverse of what was found in the anomaly judgment task. Consequently, the results suggested that A.B. has deficit in maintaining unintegrated semantic representations and M.W. has deficit in maintaining unintegrated syntactic structures. In addition, the fact that M.W.'s memory span performance was normal suggested that the capacity for retaining syntactic structure retention ability was not tapped by those span tasks. These results provided strong evidence for the separation between phonological, syntactic and semantic retention capacities, with the lexical-semantic and syntactic capacity involved in verbatim recall.

In a recent study, Martin and He (2004) provided further evidence for the multiple-capacities view by examining a new patient M.L., and comparing him to the results for EA and AB in the Martin and Romani (1994) study. M.L. performed like A.B. in showing a semantic STM deficit on span tasks. In addition to the sentence anomaly task and grammaticality judgments task used in Martin and Romani's experiment (1994), this study included an attribute questions task in which the patient has to respond with the correct noun to a simple question, e.g. "which is soft,

sandpaper or cotton?" The results confirmed the previous finding that contrasting short-term memory deficits could predict the patients' performance on the sentence-processing tasks: patient M.L. performed poorly on the sentence anomaly task when several adjectives or nouns preceded a noun or verb attribute questions, as did A. B. Both M.L. and A.B. had great difficulty with the attribute judgment task. In contrast, the patient with a phonological retention deficit (E.A.) performed normally on the attribute judgments. Martin and colleagues suggested that such results revealed that certain components (semantic STM) of the memory processes system, which is a multiple-capacity working memory system, is involved in sentence processing. According to this account, simple span tasks like digit span tap mainly the phonological component whereas the widely used complex span tasks (e.g. reading span, operation span) tap both phonological and semantic components of working memory. But only the semantic component is involved in sentence comprehension in cases in which several word meanings must be maintained before being integrated with earlier parts of the sentences.

A commonality of Caplan et al. (Caplan & Waters, 1999) and Martin et al.'s (Martin & Romani, 1994) approaches is that they both propose that there is no relationship between the retention of specifically syntactic structural information and other types of verbal information. The approach of Martin et al. claims, however, that there is a relationship between semantic capacity tapped by span tasks and the semantic capacity involved in sentence processing. Thus, a primary motivation of the present study was to determine whether Martin and colleagues findings with patients could be extended to a normal population to determine whether performance on a semantic STM task such as category probe would determine performance on semantic but not syntactic aspects of sentence processing. Moreover the use of normal subjects

allowed for the investigation of this question using sentence materials beyond the list –like structures employed with patients.

In general, there is ample evidence in favor of the storage-based account, according to which individual differences in sentence comprehension should be contributed to limited working memory capacities, although researchers have not come to a final conclusion about whether linguistic processing relies on a shared pool of working memory resources as other verbal working memory tasks, or a specific linguistic processing resource, or only certain components within working memory system. An important implication from Gordon et al. (2002) and Fedorenko et al.'s (2006, 2007) studies is that they suggested that the WM capacity underlying language processing should be conceptualized as related to the *degree of interference* rather than number of items that could be held during sentence processing. As discussed below, the role of interference has been emphasized in cue-based retrieval accounts of working memory in sentence processing.

1.3 Retrieval-based Account

In contrast to the debates about domain-general WM or language-specific WM resource, the retrieval-based account has suggested that individual differences in sentence comprehension derive from variation in the efficacy of retrieval mechanisms instead of the storage capacity of WM because number of items that can actively be maintained in the focus of attention is extremely limited and should be within everyone's capacity. Cowan (2000) summarized a wide variety of data on short-term memory capacity limits and suggested that people could only maintain four chunks in the focus of attention, which is smaller than the "magic number seven" proposed by Miller (1956). McElree (2006) reviewed the studies on the speed and accuracy of retrieving representations of recently processed information from memory and

suggested that there were two different representational states for this information. There was a clear distinction between information stored within the focus of attention and information passively stored in memory, but not a further distinction between WM and LTM. The information in the focus of attention affords privileged access, while the information outside of attention must be retrieved into the focus of attention for processing. In many tasks, the information maintained in focal attention will typically be the last item or chunk of information prior to test. The corresponding figure of this bipartite architecture is shown in Figure 1.

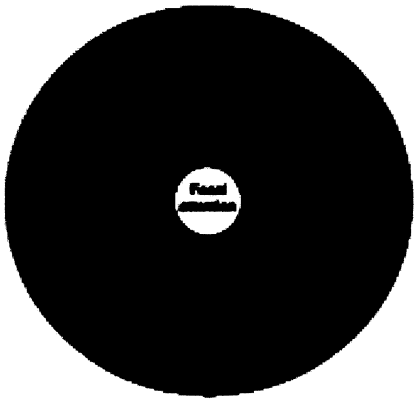


Figure 1. Bipartite architecture proposed by McElree (2006).

Some other researchers have adopted similar views of WM for sentence processing by suggesting that there are only 1-2 items that can be actively maintained in focus of attention during sentence processing, which is much smaller than that revealed by memory span measures and which should be within everyone's capacity (Lewis, 1996; McElree, 2003). What is more, previous findings revealed that people are doing parallel retrieval that involves a match of cues against all items in memory (McElree, 2006). McElree (2006) reviewed evidence from speed and accuracy tradeoff (SAT) paradigms that the estimated parallel memory retrieval time is about 80-90ms, which should be fast enough to support sentence comprehension. On the other hand, there was also compelling evidence from SAT paradigms which showed

that access to serial order information is sequential and too slow to support sentence comprehension. Thus McElree concluded that serial order information is not being retrieved during sentence processing. There are several theoretical frameworks consistent with this severe constraints hypothesis and try to explain how sentence processor works with the capacity of only two items with no serial order. Some researchers advocated a cue-based parsing mechanism to account for how the sentence processor works with a capacity of only two items and without relying on serial order information. In this approach, the functional requirements for WM are met by a combination of a limited focus attention and rapidly encoding and retrieving information from secondary store (Lewis & Vasishth, 2005; McElree, et al., 2003; Van Dyke & Lewis, 2003). Furthermore, Lewis, Vasishth and Van Dyke (2006) suggested that individual differences might be manifested in different abilities to resolve interference, instead of the storage capacity of WM.

In order to better explain the role retrieval plays in sentence comprehension, Lewis, Vasishth, and Van Dyke (2006) have proposed a different theoretical framework for the WM system that incorporates several independently motivated principles of memory. The computational principles of this framework include: 1) the focus of attention during sentence comprehension is extremely limited; 2) people are doing content-based parallel retrieval when parsing sentence, and accessing serial order information is too slow to be recruited in real-time sentence processing; 3) there are similarity-based interference effects during both encoding and retrieval processes. Lewis and colleagues suggested that this framework could also account for the storage effects found in previous studies: according to this retrieval-based model, there is only passive storage and decay, thus no distinct memory cost was caused by storage. However, difficulty in sentence processing occurred when retrieval processes

suffered from the proactive or retroactive interference.

Interference Effects and Sentence Comprehension

According to the retrieval-based account, an important concept is *interference* during retrieval processes. During sentence comprehension, the processing of each word generates cues that are used to access earlier information to be linked with the current word. The retrieval cues are a subset of the features of the item to be retrieved, and they are derived from the incoming word, context, and grammatical knowledge (Lewis, 1996; Lewis, Vasishth & Van Dyke, 2006). Interference is hypothesized to occur when these cues partially match the features of non-target information in the sentence. For example, for the center-embedded sentence "*The client who implied that the visitor was important was waiting in the office*", the comprehender needs to find a subject NP that could "wait" in order to comprehend the verb phrase "*was waiting*". Both the "*client*" and "*visitor*" partially satisfied the retrieval cues generated by the verb based on these semantic and syntactic features (i.e., both are subject NPs and both can wait). While only "client" has the feature of not yet having been assigned as the subject of a verb, some interference from "*visitor*" would be obtained because of the partial match to the retrieval cues. Importantly, this interference occurs even though "visitor" is syntactically unavailable since it has already been assigned as the subject of the copula ("was") in the relative clause. Interference effects can be either proactive or retroactive. Proactive interference results from previous information interfering with later information and causes difficulty in retaining new memories, while retroactive interference results from later information interfering with earlier information and causes forgetting of older material. Interference resolution refers to the process of resolving conflicts between targets and distracters. Participants need to access the target items while ignoring irrelevant or no-longer-relevant information.

Evidence supporting the retrieval account over the storage account was obtained from Van Dyke and Lewis's (2003) study. In this experiment, Van Dyke and Lewis manipulated both distance and interference in garden path sentences and unambiguous sentences. The examples of the unambiguous sentences are shown in Example 1.

a. The assistant forgot that the student *was standing* in the hallway.

b. The assistant forgot that the student who was waiting for the exam *was standing* in the hallway.

c. The assistant forgot that the student who knew that the exam was important *was standing* in the hallway.

Van Dyke and Lewis increased the distance between the subject and verb of the main clause by adding an intervening relative clause to the second NP. Compared to the sentence in Example1.a, Example1.b contains an intervening relative clause and a prepositional phrase. The contrast between Example1.b and Example1.c provided an estimate of interference effect. In the low interference condition (Example1.b), the intervening noun ("exam") is the object of a prepositional phrase. In the high interference condition (Example1.c), the intervening noun ("exam") is the subject of the sentential complement. The NP *exam* was a more similar distractor to the target "student" (both are subjects) in the high interference condition, and thus, there should be more difficulty in locating the appropriate subject NP. During the experiment, self-paced reading time for each word was recorded. According to the storage account, processing of "was standing" should be more difficult in sentence Example1.b than Example1.a. However, no distance effect was found on either reading times at the crucial region or in grammaticality judgment accuracies. In contrast, the results revealed that there was an interference effect on both reading time and accuracies with longer reading times and lower accuracy for sentence Example1.c than either Example1.a or Example1.b. These results thus supported the retrieval account over

the storage account. The findings are consistent with the assumption that there is no distinct memory cost associated with storage per se.

Such effects are beginning to be recognized, but the conditions causing these effects are still not well understood (Van Dyke, 2007). Several hypotheses have been proposed to account for interference effects. Gordon and colleagues suggested that interference effects are due to sentence NPs sharing similar referential characteristics. In addition to the experiment from Gordon et al. (2002) reviewed earlier on the effects of an extraneous load, Gordon, Hendrick, and Johnson (2001) provided another piece of evidence for this hypothesis by manipulating the similarity of referential characteristics between the first and second NP (common NPs, pronouns or proper names) in subject- and object-extracted relative clause sentences. Sample stimuli are shown in Example 2.

Example 2.

a. The barber that the lawyer/you admired climbed the mountain.

b. The barber that admired the lawyer/you climbed the mountain.

Subjects were required to do word-by-word, self-paced reading. The classical disadvantage of object-relative clauses over subject-relative clauses was reduced when the two nouns were of different referential types (e.g., one common noun paired with a pronoun/ proper name) compared to the situation where both were of the same type (e.g., two common nouns). Gordon et al. attributed these results to a reduction in similarity-based interference: when the two NPs had similar memory representation, such similarity could cause interference in retrieving information.

However, some researchers have argued against similarity-based interference account by suggesting that similarity of words is not the important factor; rather, it is the degree of match to retrieval cues that affects sentence comprehension difficulty. Supporting evidence was obtained from a study by Van Dyke and McElree (2006).

They adopted an extraneous memory load paradigm as in Gordon et al.'s experiment, but manipulated the semantic relation between the words in the memory list and words in the sentence. Examples are shown in Example 3. In sentence 3.a, the words in the memory list could not serve as the objects of *sailed* in the sentence, while in sentence 3.b, they could serve as objects of *fixed.* The results revealed that reading times on the verb increased significantly for sentences like that in Example3.b where the nouns matched the retrieval cues for the verb. Van Dyke and McElree argued that these results provided strong support for the cue-based parsing approach which implicates the partial match between retrieval cues and distractor items as causing parsing difficulty (e.g., Van Dyke & Lewis, 2003; Van Dyke, 2007).

Example 3.

a. Memory list: table-sink-truck
   It was the boat that the guy who lived by the sea sailed in two sunny days.

b. Memory list: table-sink-truck
   It was the boat that the guy who lived by the sea fixed in two sunny days.

There are some alternative theories accounting for retrieval difficulty in comprehending complex sentences. Traxler, Morris and Seely (2001) found that the difficulty associated with object-relative clauses was greatly reduced when the sentential subject was inanimate (e.g., "The movie that the director admired was funny"). They argued that the processing difficulty for object relatives is caused by the comprehenders' *initial assumption* that the sentential subject is the subject of the relative clause as well as the main clause. When subjects realized their error after processing the embedded subject and verb, they had to revise this assumption. Traxler et al. argued that the revision of this assumption is more difficult when the sentential subject is animate comparing to the when it is inanimate since animate objects are typically subjects whereas inanimate objects are not. Thus, Traxler and colleagues

suggested that it is the *animacy* of the nouns in the matrix and relative clause subject positions that causes the processing difficulty, not increasing working memory load or retrieval difficulty. In an fMRI study, Chen and colleagues (2006) found evidence consistent with Traxler et al.'s findings in that greater activation in the left inferior frontal gyrus for object than subject relative clause forms was only found for sentences with an animate head noun.

However, it should be noted that cue-based retrieval could account for these findings if there is difficulty in locating the *subject* of the embedded verb. That is, if the parser is searching for this subject on the basis of semantic as well as syntactic features, then *animacy* might be a semantic feature that is part of the set of cues that are searched for when the verb requires (or prefers) an animate subject. Thus, a greater degree of match to features will occur when the head noun is animate than inanimate.

Additional evidence supporting a cue-based retrieval account for interference resolution during sentence processing has been obtained by Van Dyke (2007) in a further study. Van Dyke manipulated both semantic interference and syntactic interference in one experiment. Example sentences from her study are shown in below.

Example 4.

a. LOW Syntactic and LOW Semantic interference condition

The client who had arrived after the important meeting that day was waiting in the office.

b. LOW Syntactic and HIGH Semantic interference condition

The client who had arrived after the important visitor that day was waiting in the office.

c. HIGH Syntactic and LOW Semantic interference condition

The client who implied that the meeting was important that day was waiting in the office.

d. HIGH Syntactic and HIGH Semantic interference condition

The client who implied that the visitor was important that day was waiting in the office.

Semantic interference is caused by the match between the intervening NP and semantic retrieval cues appropriate to the verb. For example, for sentences 4(b) and

4(d) in Example 4, the intervening NP *visitor* fits the semantic cues of the verb *was waiting* while the intervening NP *meeting* in sentences 4(a) and 4(c) do not. There were two main findings from this experiment. First, Van Dyke observed both semantic and syntactic interference effects. Semantic interference occurred even when the syntactic properties were inappropriate (that is, the interfering noun could not serve as the subject of the verb). Second, the time course of these two interference effects was different: the syntactic interference effect occurred earlier than the semantic interference effect. The finding that NPs whose syntactic character is inappropriate are nevertheless recruited when they are semantically appropriate is crucial to the relationship between language processing and the more general mechanisms of memory. That is, this finding presents a challenge for purely grammar-driven parsers and highlights the importance of the properties of memory retrieval mechanism, which are involved in resolving grammatical dependencies.

Individual differences in retrieved-based account

Since cue-based retrieval interference effects have been confirmed during sentence processing, one might hypothesize that individual differences in retrieval abilities would relate to the degree of interference effects. Van Dyke also suggested "the broader implications of this research indicate a need for a precise specification of the interface between the parsing mechanism and the memory system that supports language comprehension". Motivated by the results from these studies, we were interested in looking at the semantic and syntactic interference effects in sentence comprehension and their relationship to working memory capacity. In the current study, we used the same materials as in Van Dyke's (2007) experiment. To investigate the possible relations between working memory capacity and sentence comprehension

as claimed by different accounts, we employed a set of WM measures (operation span, reading span, digit span and category probe), an executive function task (standard Stroop test), and some verbal ability tests (verbal SAT, WAIS) in our experiment. Based on the results from previous studies (Lewis, Vasishth & Van Dyke, 2006; Van Dyke, 2007), we expected to find both syntactic interference and semantic interference effects in current experiment. What is more, we hypothesized that the syntactic interference effects will occur earlier than the semantic interference effects as in Van Dyke's experiment. Van Dyke discussed two possible reasons for this time course difference. First, it is possible that it took longer for the subjects to compute the semantic association between the distractor and the verb (McElree & Griffith, 1995, 1998, cited in Van Dyke, 2007). Second, the semantic interference effects may be part of sentence wrap-up processing, which would not occur in the critical region (Van Dyke, 2007).

There are several possible outcomes for the relations between measures of span and interference effects. The predictions are shown in Table 1, and as shown in the table: 1) according to Martin et al.'s multiple capacities account (1994, 2003, 2004) in which the semantic component but not the phonological component is assumed to be related to sentence processing, the semantic STM measure should correlate with semantic interference resolution, while the phonological STM component should be uncorrelated with both types of interference. Also, a correlation with syntactic interference would not be expected because the syntactic structure retention ability is independent from semantic and phonological retention abilities; 2) according to Just and Carpenter's (1992) or Gibson et al.'s (1996, 2008) domain general resources account, there should be a relation between the size of interference effects and complex span measures, e.g. operation span and reading span because the

WM capacity underlying language comprehension is a domain general resource that is tapped by span tasks. Thus, high-span subjects should be able to maintain more accurate representations of more items in WM leading to a greater ability to discriminate targets and distractors; 3) if there are only correlations between WM span tasks and interference resolution in offline sentence processing (accuracy or RT to comprehension question), but not online sentence processing (self-paced RT), such will provide supporting evidence to Caplan and Water's domain specific resources account (1999) that sentence processing makes demands on WM that not tapped by span tasks – supporting the notion of a specific capacity for syntactic processing; 4) based on Van Dyke and McElree's retrieval-based model (2006), we expected to see no relation between interference effects and WM or STM measures, but we might obtain a relation between ability to resolve interference as tapped by the Stroop task if interference resolution processes are common across different types of tasks.

Table 1. Predictions of the relations between individual differences measurements and interference effects during sentence processing.

Account	Predictions
Multiple capacities	The semantic STM measure should correlate with semantic interference resolution.
	The phonological STM component should be uncorrelated with both types of interference
	Syntactic interference resolution ability should not be correlated with either semantic STM or phonological STM.
Domain general resources	Complex span measures (e.g. reading span and operation span) should be correlated with the size of all interference effects.
Domain specific resources	There should be correlations between WM span tasks and offline interference effects only, but not with online interference effects.
Retrieval-based model	There should be no correlation between interference effects and WM or STM measures once language knowledge is controlled for.
	Performance in Stroop task should be correlated with the size of interference effects.

2 Experiment

2.1 Pilot Study

2.1.1 Method

2.1.2 Subjects

Fifty-six undergraduate students from Rice University were recruited for this experiment. All the subjects were native English speakers without a diagnosed reading or learning disability, and had normal or corrected-to-normal vision. Subjects got credit toward experiment participation requirements for their courses.

2.1.3 Materials, design, and procedure.

Reading comprehension task

For our materials, we modified the sentences used in Van Dyke's experiment (2007) by adding a phrase to the last region to avoid the confounding of "sentence wrap-up" effects (Just, Carpenter, & Woolley, 1982; King & Just, 1991). The sentences are shown in Appendix A. There were 48 sets of sentences with four different types of sentences in each set, as shown in Table 2. For each set of sentences, all four sentences began with the same introduction region and differed in the intervening region, in which semantic and syntactic interference were manipulated, which resulted in four versions of each sentence. To increase readability, we refer to the low and high syntactic interference conditions as LoSyn and HiSyn, while the low and high semantic interference conditions are called LoSem and Hisem. After the intervening region, the main verb for the long-distance dependency was identified as the critical region. The first phrase following the main verb was identified as the spillover region, because it is very likely that the effect will spill over to the next region after the critical region (Just, Carpenter, & Woolley, 1982). That is, unfinished processing of the previous phrase often has some influence on the time spent on the

following phrase. The rest of the sentence was defined as the final region, on which we might obtain some sentence wrap-up effects (Just, Carpenter & Woolley, 1980; King & Just, 1991).

Table 2. Example Syntactic and Semantic Interference Stimuli for Experiment Showing Phrasal Regions for Self-Paced Reading

Sentence region		Example stimulus
introduction		The worker was surprised that the resident
Intervening region	LoSyn/LoSem	who was living near the dangerous warehouse
	LoSyn/HiSem	who was living near the dangerous neighbor
	HiSyn/LoSem	who said that the warehouse was dangerous
	HiSyn/HiSem	who said that the neighbor was dangerous
Critical region		was complaining
Spillover region		about the investigation
Final region		in the morning.

Note. "Lo-" and "Hi-" refer to low and high interference condition, while "-Syn" and "-Sem" refer to syntactic interference and semantic interference condition.

Each subject saw eight target sentences in each condition. To avoid repetition of the verbs and sentence content within one participant, the four items in each set were assigned to four lists and each subject received only one sentence from each set. The selection of sentences from each set that appeared in each condition was counterbalanced across the four lists. Two pseudo-randomized sequences were created for each stimulus list, resulting in a total of eight lists. Each participant received one of the eight lists in the experiment. Additionally, ninety-six filler items were also constructed and presented to each subject. Half of the fillers were short grammatical sentences and the other half of the fillers was long grammatical sentences. Moreover, half of the long sentences contained embedded object relative

clauses, which increased the processing difficulty of these sentences and may have distracted attention away from the actual target sentences. Thus the number of sentences in each condition presented to each participant was 48 experimental sentences and 96 fillers for a total of 144 sentences. The pseudo-randomized order in each version had the following constraints: a) the beginning three sentences of each block are fillers; b) the same experimental sentence type did not repeat in three successive sentences. After the presentation of the sentences, there was a comprehension question for each experimental sentence and for half of the fillers to encourage participants to integrate incoming material into a consistent interpretation.

During the experiment, stimuli were presented in a phrase-by-phrase, non-cumulative and self-paced fashion (Just, Carpenter, & Woolley, 1982) on a Macintosh computer with PsyScope. Before the experimental sentences, there were 12 practice sentences to help the participants get familiar with the experiment. Participants were instructed to read each sentence for comprehension and told that there would be a comprehension question after some sentences. All trials began with a fixation point appearing in the center of the screen beginning for 1000 ms, and then followed by the first phrase. The participants had to press a button with their index finger to bring up the phrases in each sentence. A period was presented together with the last phrase. If the sentence was followed by a comprehension question, the subjects were instructed to indicate the correct answer by pressing button as quickly and accurately as possible. For instance, the question for the example above was "*Who was complaining*? " and the two alternative answers were *resident* or *worker/neighbor*. Subjects pressed one button for the left answer using the index finger, and another button for the right answer using the middle finger. All the experimental sentences and half of the fillers were followed by a comprehension question. The questions for the fillers tapped

positions rather than the second noun (e.g. *resident*), which is the correct answer to the comprehension question in critical sentences. The position of the answers was counter-balanced such that half of the questions had the correct answer on the left side and half had the correct answer on the right side. The question sentence remained on the screen until either the subjects responded or 2s had elapsed. The next sentence started after an inter-trial interval of 1000ms. The reading time for each phrase, accuracy for comprehension questions, and speed of question answering were recorded for each trial.

Operation span task

We used the automated version of Operation Span (Aospan) to measure working memory capacity. There were three practice sections to help subjects get familiar with the task: (a) the first practice section was a simple letter span task. Letters appeared sequentially on the screen, and the subjects recalled the letters in the same order in which they were presented by clicking the box next to the appropriate letters (no verbal response was required). (b) in the second practice section, a math operation (e.g., (1*2)+1=?) was presented on the screen. The subjects were instructed to solve the math operation as soon as possible, and click the mouse to see the next screen. Then there was a number presented on the screen and the subjects should click either on either the "true" or "false" box by comparing this number to their answer. (c) The third section contained both letter recall and math portions in the same format as in the experimental trials. A math operation was presented on the screen first. After subjects solved it and clicked the mouse, a digit was presented for judging whether it was the correct answer. Then a letter to be recalled was shown on the screen for 800 ms. After the end of each set, the subjects have to recall the letters in the correct order using the method described in a). The computer calculated each individual's mean

time of solving the equations during the second practice section, and this time was used as a time limit for the math portion of the experimental session. After participants completed three practice sections each of set size 2, the program progressed to the experimental trials. The experimental trials contain three sets at each set size, with set sizes ranging from 3 to 7 items. This results in a total of 75 sets with 75 letters and 75 math problems. The order of set sizes was random for each participant. We evaluated the subjects' performance by two scores reported by the system: Operation span score (use traditional absolute scoring method, it is the sum of the perfectly recalled sets) and total number correct (total number of letters recalled in the correct position).

Category probe and Rhyme probe tasks

In the Category probe task, subjects were presented with an auditory word list. After a short pause, they heard a probe word and had to judge whether this word was in the same category as any of the words in the list. (All the words in one list were from different categories.) The subjects were shown a list of all the words that would be presented in the experiment, as well as the categories to which they belonged. The number of words in each list ranges from 4 to 7. There were some practice trials before each level to help the subjects get familiar with the task. There were 24 lists at each list length. We collected the overall accuracy for each subject.

In the rhyme probe task, subjects judged whether the probe word rhymed with any of the list words. Because we got a ceiling effect when using the patient version of the rhyme probe task (e.g., Martin et al., 1994), we increased the difficulty of this task by changing all the distractors to be phonologically related to the probe word. For both the no-trials and yes-trials, each distractor overlapped with the probe on either the initial consonant or consonant cluster (e.g., st, tr, sl, etc), the vowel, or the final

consonant (or cluster). The number of words in each list ranges from 4 to 7, and there are 24 lists in each level. There were practice trials as well. We collected the overall accuracy for each subject.

Stroop task

We adopted the classical Stroop task in the current experiment. Subjects were required to name the ink color in all conditions. There were three conditions: (a) congruent, in which a color word appeared which was congruent with the ink color; (b) incongruent, in which a color word was presented in a different ink color (e.g., the word blue written in red); (c) neutral, in which the subjects were presented a series of colored asterisks. Response naming latencies were recorded from the onset of the stimulus. We computed the Stroop interference score for each subject by subtracting the average reaction time in the neutral condition from the incongruent condition.

Digit span task

In the digit span task, the subjects heard a list of digits. And they were required to repeat the numbers aloud in order after the list was finished. The number of digits in each list ranged from 3 to 9, and there were 2 trials at each level. We computed the overall accuracy for each subject.

2.1.4 Results

The sentence comprehension experiment produced three dependent variables: reading times (RT) from self-paced reading and reaction time and accuracy for the comprehension questions. For reading time, only data from trials in which participants answered the question correctly were included in the analysis. Outliers were calculated by condition for each subject and reading times above or below 2.5 standard deviations from the mean for each condition were removed. As mentioned above in the materials section, there were four regions of interest: critical region,

spillover region, last phrase, and the comprehension question. The mean accuracy and reading time for each of these four regions are shown in Table 3 and Figure 2. As expected, the subjects showed the highest accuracy and shortest reading time in the LoSyn/LoSem condition in all the regions. However, for the reading time data, the results revealed that the subject showed longest RT in the LoSyn/HiSem condition in most of the regions except the spillover region, rather than in the HiSyn/HiSem condition as we expected.

Table 3. Mean Accuracy and Reading Times in Sentence Comprehension task (ms)

		Reading time (ms)			
Interference type	Accuracy	Critical	Spillover	Last phrase	Question
LoSyn/LoSem	.86	794	774	725	2224
LoSyn/HiSem	.83	890	827	785	2765
HiSyn/LoSem	.84	874	836	738	2312
HiSyn/HiSem	.82	844	834	744	2632

Note. LoSyn and HiSyn refer to low and high syntactic interference conditions. And LoSem and HiSem refer to low and high semantic interference conditions.

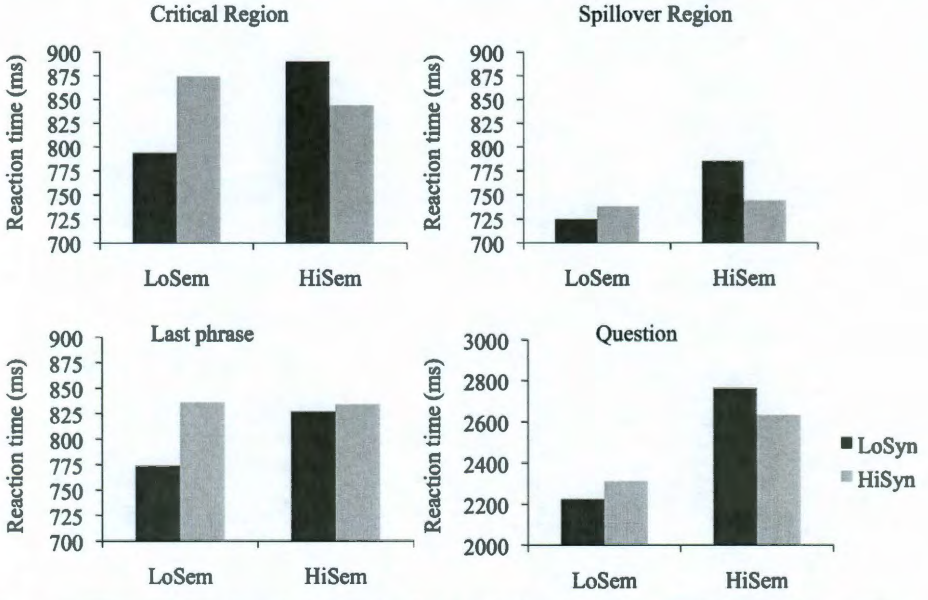


Figure 2. The reading time (ms) on the four regions of interests: critical region, spillover region, last phrase, and comprehension question.

Main effect Analysis

Both accuracy and reading time measures were analyzed via a 2 (high or low semantic interference) × 2 (high or low syntactic interference) factorial repeated measures ANOVA using error terms based on participant (F_1)[1]. The results are shown in Table 4.

For the accuracy of question answering, there was a marginal effect of semantic interference with the high interference condition being more difficult than the low interference condition (Low interference, 0.83; High interference, 0.85), F (1, 55) = 3.11, p = .08. Similarly, analysis of RTs for the comprehension question showed a main effect of semantic interference (Low interference, 2268 ms; High interference, 2698 ms), F (1, 55) = 42.03, p < .001. The main effect of syntactic interference and the interaction between semantic and syntactic interference were not significant in either of these two analyses.

In the analysis of reading times in the critical and spillover regions, we expected to obtain main effects of semantic and syntactic interference and perhaps an interaction such that the semantic interference effect was greater in the high syntactic condition and syntactic interference was greater in the high semantic interference condition. However, this is not the pattern that was obtained as is evident in Table 3 and Figure 1. The analysis of reading times in the critical region showed that the main effects of semantic interference and syntactic interference were not significant. A significant interaction was observed, F (1, 55) = 5.17, p= .03, but the interaction was the opposite that predicted. That is, the semantic interference effect was significant in the LoSyn condition (Low semantic interference condition, 794 ms;

[1] For the pilot experiment, we only reported F_1 analysis but did not include F_2 analysis, because we decided to redesign the sentences for a subsequent experiment. See more details in discussion section.

high semantic interference condition, 890 ms), F (1, 55) = 7.52, p = .008, MSE = 34,820, but not in the HiSyn condition (Low semantic interference condition, 890 ms; high semantic interference condition, 844 ms), $F < 1$. The syntactic interference effect was significant in the LoSem condition (Low syntactic interference condition, 794 ms; high syntactic interference condition, 874 ms), F (1, 55) = 4.39, p= .04, MSE = 41,456, but not in the HiSem condition (Low syntactic interference condition, 890 ms; high syntactic interference condition, 844 ms), p=0.20. (None of the effects reached significance in the spillover region and last phrase.)

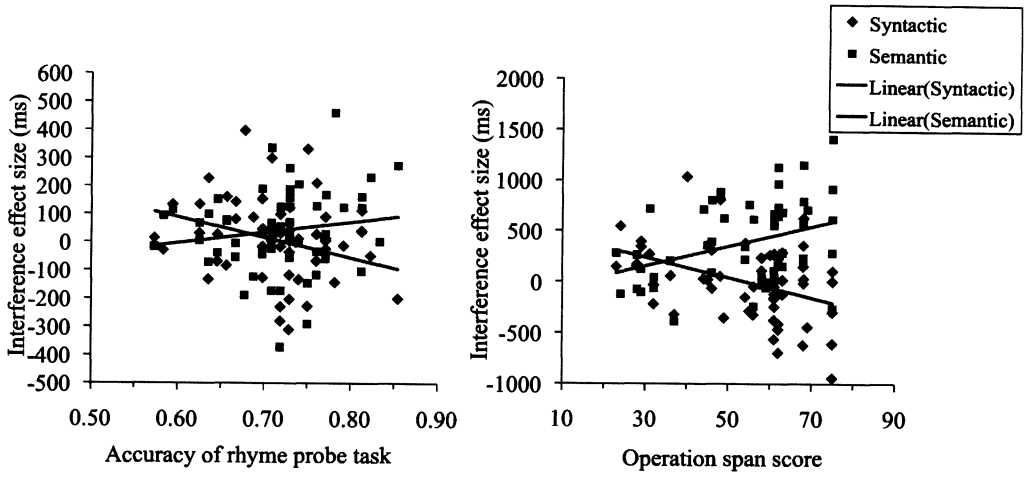
The RT data showed the unexpected result that the longest times for the critical region, the spillover region, and the question were found for the low syntactic, high semantic condition with the times for the high syntactic, high semantic condition being somewhat faster (though non-significantly so). The possible source of this unexpected pattern will be addressed in the discussion section.

Table 4. Analysis of Variance for all Dependent Measures

	Main effect		
Measure	Syntactic interference	Semantic interference	Interaction
Accuracy	$F(1,55) = 1.96$, p=.17, $MSE = 0.012$	$F(1,55) = 3.11$, p=.08, $MSE = 0.012$	$F(1, 55) = .17$, p=.68
Reading time			
Critical	$F(1,55) <1$, ns	$F(1,55) = 2.36$, p=.13, MSE = 26,151	$F(1,55) = 5.17$, p=.03, MSE = 43.680
Spillover	$F(1,55) = 1.64$, p=.21, MSE = 65,390	$F(1,55) <1$, ns	$F(1,55) = 2.51$, p=.12, MSE = 17,436
Final	$F(1,55) <1$, ns	$F(1, 55) = 2.38$, p=.13, MSE = 25,420	$F(1,55) = 2.88$, p=.095, MSE = 14,465
Question	$F(1,55) <1$, ns	$F(1,55) = 42.03$, p<.001, MSE = 243,293	$F(1,55) = 2.64$, p=.11, MSE = 256,938

Correlation Analysis

To test the correlations between individual differences measures and interference resolution, we calculated interference effect size for both syntactic interference and semantic interference by using averaged RT or accuracy in high interference condition minus the low interference condition. We correlated the semantic interference effect size and syntactic interference effect size with the performance on other tasks (Category probe task, Rhyme probe task, Operation span, Digit span, and standard Stroop task) for all the regions. The results showed that there was a significant negative correlation between rhyme probe test and degree of syntactic interference effect in the critical region, r = -0.281, p = .032, which suggested that the better the subjects' phonological short-term memory, the less syntactic interference effect they showed. The correlation between semantic interference effect in the critical region and rhyme probe performance was not significant (p = .19). In addition, in the question region, there was a negative correlation between operation score and syntactic interference effect size (r = -.387, p = .004), but a positive relation between operation span and semantic interference (r = .356, p = .008). None of other correlations between interference effects and WM tasks or Stroop task reached significance. The scatter plots of these significant correlations are shown in Figure 3.



A. Critical Region B. Question Region

Figure 3. Correlations between interference effect size and working memory capacity.

Figure 3(A) displayed the correlation between accuracy of rhyme probe task and interference effect size in critical region. Figure 3(B) displayed the correlation between operation span score and interference effect size during question answering. In both figures, the blue dots and line represent syntactic interference effect. The red dots and line represent semantic interference condition.

2.1.5 Discussion

The motivation of our study was to investigate the link between working memory capacity and language comprehension ability. In the pilot study, we examined performance on a sentence comprehension task, a series of working memory tasks (Operation Span, Category Probe test, Rhyme Probe test, and Digit Span) and an executive function task (standard Stroop task). Our results partly replicated Van Dyke's (2007) results which showed both syntactic and semantic interference effects. However, we did not find a time course difference between these two interference effects. Moreover, we obtained an anomalous effect in which the most difficult condition was the low syntactic/high semantic condition rather than the high syntactic/high semantic condition. Consequently, the syntactic interference effect was only evident in the low semantic interference condition and the semantic interference effect was only evident in the low syntactic interference condition.

Neither effect spilled over to the next region after the critical region.

There are several possibilities that might contribute to the differences between our results and Van Dyke's (2007). There were several methodological differences in that we used phrase-by-phrase reading and she used either word-by-word reading or eye-tracking. Also, the nature of the comprehension questions was different.

What appears to be the case is that with the results are showing a *local coherence effect* which complicates the interpretation of the semantic and syntactic interference effects. The local coherence effect refers to the phenomena that subject compute partial parses that are syntactically compatible with only a proper subpart of the input (Tabor et al., 2004). For example, in the sentence "The mechanic who maintained the truck was working hard", subjects may experience difficulty when processing the phrase *was working* because the *truck* is near "was working" and is semantically and syntactically compatible as the subject of "was working". Even assuming a very restricted working memory capacity, it is possible that "truck" and "was working" are available simultaneously in the focus attention leading to a tendency to grammatically link the two. In our experiment, in the low syntactic/high semantic condition, the interfering noun appears right before the critical verb and thus there may be a strong tendency to assume this noun is the subject. Thus, the unexpectedly long times in this condition might have been due to a contribution of a local coherence effect. Some previous studies have suggested that it is necessary to take local coherence effects into consideration when determining if there is an interference effect (Gibson, 1998; Tabor, Galantucc, & Richardson, 2004).

Van Dyke's studies (2007) addressed the possible contribution of local coherence to her semantic interference effects. That is, in her first two experiments, in the low syntactic interference conditions, the interfering noun appeared

immediately before the critical verb. The effect of semantic interference might be due to linking the semantically plausible preceding noun with the following verb. In her third experiment, Van Dyke added an adverbial phrase between the object noun phrase and the critical verb (e.g. The pilot remembered that the lady who was sitting near the smelly man *yesterday afternoon* moaned about a refund.) to break the local coherence effect. She still obtained semantic and syntactic interference even without a contribution of local coherence.

Although Van Dyke's (2007) experiments did not appear to show an effect of local coherence, it is possible that ours did, perhaps because of differences in methodology. Thus, in the next experiment, we decided to add an adverbial phrase prior to the critical verb as Van Dyke did in her third experiment in order to avoid inducing a local coherence effect.

Second, the only significant correlation we obtained for the sentence reading times was that the degree of syntactic interference effect negatively related with the performance in the rhyme probe test, which indicated that the better the phonological short-term memory, the less the comprehender was affected by the high syntactic interference condition. This result was also unexpected according to previous patient studies, which demonstrated that only patients with semantic short-term memory deficits had difficulty in sentence comprehension, whereas patients with phonological short-term memory deficits had well preserved sentence comprehension ability (Martin & He, 2004). One potential explanation was that the subjects in our experiment did not instantly integrate the upcoming phrase into the earlier part of the sentence. The phrase-by-phrase reading may have emphasized phonological storage and caused subjects to go back and verify what had been presented in earlier phrase. However, as discussed above, the interference effects we observed here could be a

mixture of interference effects and local coherence effects. Such a mixture would make interpretation of the correlations problematic.

In all, the data from the pilot experiment gave some support to some of our predictions. We found both syntactic interference and semantic interference effects. However, we failed to observe the time course difference between these two interference effects reported earlier. But what we need to take into consideration is that there might be a mixture of interference effects and local coherence effects in this experiment. Given the pervasiveness of interference effects in language processing, and its importance in understanding the direct link between working memory ability and sentence comprehension, further studies are needed. Thus in the revised experiment, we added an adverbial phrase between the intervening region and the critical region to diminish the local coherence effect.

2.2 Experiment

Based on the pilot study, we modified our sentence materials to avoid the local coherence effect by adding an adverbial phrase before the critical verb.

2.2.1 Subjects

We collected data from ninety-six subjects. All the subjects were native English speakers without a diagnosed reading or learning disability, and had normal or corrected-to-normal vision. Subjects got credit toward experiment participation requirements for their courses.

2.2.2 Materials, design and procedure

Sentence comprehension task

For the sentence comprehension task, we ran it in the same fashion as in the pilot study: subjects were instructed to do self-paced, noncumulative, phrase-by-phrase reading by pressing the button. Based on the results from pilot study, we modified our experimental sentences by adding an adverbial phrase positioned prior to the critical verb (e.g. *The worker was surprised that the resident who was living near the dangerous warehouse last month was complaining about the investigation*). In addition, instead of providing two alternatives to the comprehension question, we only presented the comprehension question (e.g., Who was complaining?) and asked subjects to provide a spoken response. RT for the question was measured through a voice key response.

We ran a pilot study on 16 subjects to see whether using this type of sentence would give promising results, but question answering accuracy was very low (mean = 0.70). Thus we decided to cut off the first noun phrase in the sentence to make it simpler. Examples of the sentences we used in this experiment are shown in Table 5. All the experimental sentences we used are shown in Appendix B. There were 80 sets

of experimental sentences. We separated the four sentences in each set into different list to construct 4 lists with 80 critical sentences and 80 fillers. Two pseudo-random lists were constructed for each list following the same constraints as in pilot study. Thus, there were 8 versions of sentences.

Table 5. Experimental materials for Experiment

Sentence region		Example stimulus
Introduction		The resident
Intervening region	LoSyn/LoSem	who was living near the dangerous warehouse
	LoSyn/HiSem	who was living near the dangerous neighbor
	HiSyn/LoSem	who said that the warehouse was dangerous
	HiSyn/HiSem	who said that the neighbor was dangerous
Adverbial phrase		last month
Critical region		had complained
Spillover region		about the investigation.

Working memory & Executive function tasks

For the working memory and executive function tasks, we kept all the tasks used in the pilot study (Operation Span, Category Probe, Digit Span, Stroop task) except the Rhyme probe task. In the pilot study, the rhyme probe task was fairly highly correlated with the digit span task, $r = 0.37$, $p = .005$. We decided to keep the digit span task only, which took much less time to complete (around 5 min) compared to the Rhyme probe task (around 20 min).

In addition, we also included the *Reading span task* in our second experiment. This is one of the most widely used tests in the studies aimed at investigating the correlation between working memory capacity and language processing (Daneman & Carpenter, 1980). Previous studies have shown that the Reading Span measure correlates highly with certain aspects of reading comprehension, such as the Verbal

SAT (Daneman & Carpenter, 1980, Daneman & Merikle, 1996). Thus, we included Reading Sapn in order to relate our findings to the substantial literature employing this measure. We used the automated version of Reading Span developed by Schrock and Engle (2005) to measure working memory capacity. The task is mouse-driven and it is very similar to the *Operation span task* as we introduced above. The difference is that rather than solving the math operation, subjects are instructed to judge whether the presented sentence make sense or not (e.g. Andy was stopped by the policeman because he crossed the yellow heaven.). After each sentence, a letter to be recalled is shown on the screen for 800 msec. At the end of each set of sentences, the subjects recalls all the letters showed in current set in order. At the beginning of the experiment, there were the same types of practice sessions as in automated version of operation span task as described in earlier section. The set size ranged from 3 to 7 items. There were a total of 75 letters and 75 sentence judgments. The order of set sizes was random for each participant. We evaluated the subjects' performance by calculating the total number of correctly recalled letter sets.

Receptive vocabulary test and verbal SAT

We also decided to add two standard measures of verbal ability. We adopted the vocabulary test from WAIS III which requires subjects to provide word definitions (e.g. Tell me what *confide* means) and we also collected the SAT verbal scores for all the subjects from the admission office of Rice University. We began the vocabulary test from the 12th item to the 33rd item in WAIS II because the words before the 12th were not discriminating enough for undergraduate students. We included the vocabulary test because there were studies reporting that receptive vocabulary was correlated with both working memory and with sentence comprehension and might be the source of some of the relation between WM and sentence comprehension

(MacDonald & Christiansen, 2002; Perfetti, 2007; Van Dyke, Clinton & Anuenue, 2010). In the study conducted by Van Dyke and colleagues, individual differences were measured by using 25 tasks, including working memory tasks. The results showed that after removing variance shared with a measure of general ability (IQ), receptive vocabulary was the only significant predictor of comprehension. Thus, Van Dyke et al. concluded that working memory capacity was not important after considering general verbal ability. Instead, only knowledge of word meanings, which could relate to the richness of semantic representations for words in the sentence, predicted comprehension. However, some studies have reported that there is a very high correlation between working memory and IQ (Hambrick & Engle, 2002). Thus, by factoring out verbal IQ, one may be factoring out working memory capacity. We wanted to investigate whether the same pattern of results would be obtained here or whether working memory measures would contribute beyond vocabulary and verbal ability. Of course, if there were no contribution from WM after factoring out IQ, there would still be an interpretive issue regarding whether WM can be distinguished from verbal IQ.

2.3 Results

2.3.1 Design and data analysis

Ninety-six undergraduate students from Rice University were recruited for this experiment. The results are shown in Table 5. Given the results from previous study (Van Dyke, 2006) and our pilot experiment, we expected to observe: 1) both syntactic and semantic interference during sentence reading and question answering; 2) a syntactic interference effect that occurs earlier than the semantic interference effect during sentence processing 3) a correlation between interference effects and WM measures.

For these sentence comprehension materials, three RT measures were of interest: reading times for the critical region and the spillover region, and time to answer the comprehension question. The reading time on the critical region and the spillover region was recorded as the time between key-presses. The reaction time for comprehension questions was recorded through a voice key response. Thus, this experiment produced four dependent measures: accuracy and reaction time for the comprehension question, as well as reading time for the two phrases. As in the pilot experiment, all the measures were analyzed via a two (high or low syntactic interference) by two (high or low semantic interference) factorial repeated measures ANOVA using error terms based on participants (F_1) and items (F_2). Only correctly answered items were included into the RT analysis.

The effect size for both syntactic interference and semantic interference was calculated by using the mean accuracy or RT in the high interference conditions minus the low interference conditions. These interference effect sizes were related to all the individual difference measures. As discussed below, the data were also analyzed using a multiple regression approach in which performance in the high

interference condition was predicted by performance in the low interference condition plus performance on the individual differences measures.

2.3.2 Results

2.3.2.1 Comprehension questions

Accuracy and reaction times for comprehension questions, and reading time results for each region are presented in Table 6. As we can see from the table, overall accuracy was very high (0.86). Subjects showed the highest accuracy in the LoSyn/LoSem interference condition (0.90), and the lowest accuracy in the HiSyn/HiSem interference condition (0.81). The results for comprehension question performance are shown in Figure 4. The error bars represented standard error of the mean (Cousineau, 2005) corrected for a within-subjects design. Since the error terms in a within-subjects design factor out variation in the overall level of performance for each subject, the scores in each condition were corrected by subtracting the subjects' mean score across conditions from it, and adding the grand mean of all the subjects (Mean (participant/each condition) – mean (participant) + mean (group)).

Table 6. Mean Accuracy and Reading Times in Sentence Comprehension task (ms)

		Reading time (ms)		
Interference type	Accuracy	Critical	Spillover	Question
LoSyn/LoSem	0.90	908	935	1286
LoSyn/HiSem	0.86	920	966	1409
HiSyn/LoSem	0.88	903	971	1301
HiSyn/HiSem	0.81	920	1019	1470

Note. LoSyn and HiSyn refer to low and high syntactic interference conditions. And LoSem and HiSem refer to low and high semantic interference conditions.

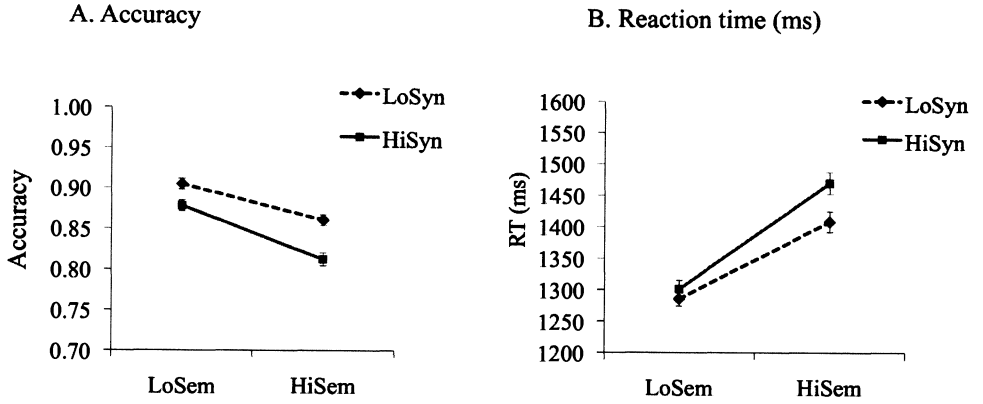


Figure 4. The overall accuracy and reaction time data for comprehension question.

The results of factorial repeated measures ANOVA in each region are shown in Table 7. The F_1 values are the results of analysis by subjects, and the F_2 values are the results of analysis by items. For comprehension question accuracy, main effects of both syntactic interference and semantic interference were observed, with the HiSyn sentences being more difficult than the LoSyn sentences (.85 vs. .88), F_1 (1, 95) = 23.68, $p < .001$, F_2 (1,79) = 16.77, $p < .001$, and HiSem sentences being more difficult than the LoSem sentences (.84 vs. .89), F_1 (1, 95) = 48.16, $p < .001$, F_2 (1,79) = 29.29, $p < .001$. The interaction was not significant, F_1 (1, 95) = 2.99, $p = .09$, F_2 (1,79) = 1.65, $p = .20$. RTs for the comprehension questions revealed a significant main effect of syntactic interference with responses in the HiSyn condition being slower than in the LoSyn condition (1386 ms vs. 1348 ms), F_1 (1, 95) = 5.83, $p = .018$, F_2 (1,79) = 8.90, $p < .001$, and a main effect of semantic interference, with responses in the HiSem condition being slower than in the LoSem condition (1440 ms vs. 1294 ms), F_1 (1, 95) = 65.64, $p < .001$, F_2 (1,79) = 65.98, $p < .001$.

Table 7. Analysis of Variance for all Dependent Measures

Measure	Main effect		Interaction
	Syntactic interference	Semantic interference	
Accuracy	F_1 (1,95) = 23.68, p< .001	F_1 (1, 95) = 48.16, p< .001	F_1 (1, 95) = 2.99, p= .09
	F_2 (1,79) = 16.77, p< .001	F_2 (1,79) = 29.29, p< .001	F_2 (1,79) = 1.65, p= .20
Reading time			
Critical	F_1 (1,95) <1, *ns*	F_1 (1, 95) =1.63, p= .82	F_1 (1, 95) <1, *ns*
	F_2 (1,79) <1, *ns*	F_2 (1,79) <1, *ns*	F_2 (1,79) <1, *ns*
Spillover	F_1 (1,95) = 13.59, p< .001	F_1 (1, 95) = 5.48, p= .02	F_1 (1, 95) <1, *ns*
	F_2 (1,79) = 5.62, p= .02	F_2 (1,79) = 8.21, p= .005	F_2 (1,79) <1, *ns*
Question	F_1 (1, 95) =5.83, p= .02	F_1 (1, 95) = 65.64, p< .001	F_1 (1, 95) = 1.69, p= .20
	F_2 (1,79) = 8.90, p< .001	F_2 (1,79) = 65.98, p< .001	F_2 (1,79) = 1.71, p= .20

2.3.2.2 Self-paced reading times

Figure 5 shows the mean reading times in each region for the four conditions. The error bars represented corrected standard error of the mean. For the analysis of sentence reading times, there was a trend toward a semantic interference effect in the expected direction (920 ms in the high interference condition vs. 905 in the low interference condition), whereas there was no sign of a syntactic interference effect (912 ms in the high interference condition and 914 ms in the low interference condition). However, all effects were far from significant (all Fs <1). A significant main effect of syntactic interference was observed in the spillover region, with responses in the HiSyn being slower than in the LoSyn conditions (995 ms vs. 950 ms), F_1 (1,95) = 13.59, p< .001, F_2 (1,79) = 5.62, p = .02. Also, there was a significant main effect of semantic interference, with HiSem condition being slower than the LoSem condition (993 ms vs. 953 ms), F_1 (1, 95) = 5.48, p = .02, F_2 (1,79) = 8.21, p = .005. The interaction was not significant, Fs<1.

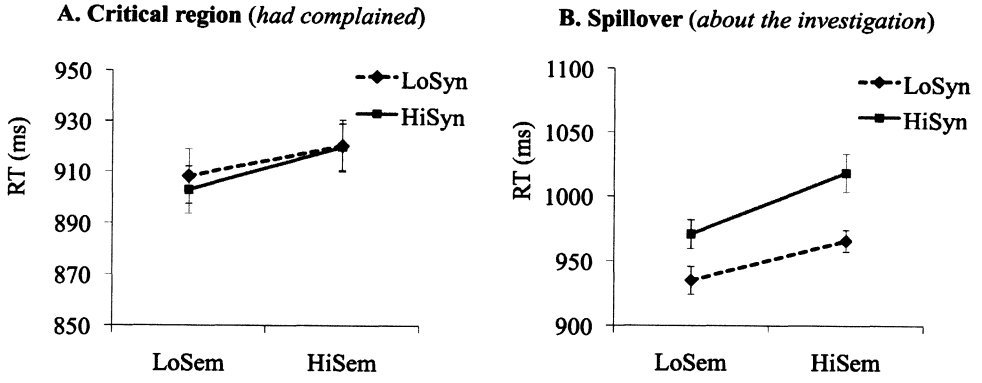


Figure 5. RTs for self-paced reading.

Sentence processing summary. The results replicated Van Dyke's (2007) experiment in showing that there were both syntactic and semantic interference effects. However, there were no time course differences for semantic vs. syntactic effects, contrasting with what was found in Van Dyke's experiment. This could be attributed to the fact that there was more power in our experiment since we had a larger sample size (96 subjects vs. 35 subjects). In the spillover region, the mean syntactic interference and semantic interference effects were of similar size, but the syntactic effect was less variable across subjects. In the question region, the semantic interference effect size was larger than the syntactic interference effect size (146 ms vs. 39 ms). Thus, results more in line with those of Van Dyke would most likely have been obtained with a smaller sample size. The local coherence effect that was found in the pilot study was eliminated here.

2.3.3 Correlational Analyses

We correlated the size of the semantic interference and syntactic interference effect with the performance on other tasks (category probe task, reading span, operation span, digit span, standard Stroop and vocabulary score from WAIS III) in all the regions. That is, for each subject, we calculated a difference score between the high and low semantic interference conditions and between the high and low syntactic interference conditions. These difference scores were correlated with our individual differences measures.

Before running correlational tests, we calculated the reliabilities for all the individual differences measures and dependent measures (different scores). One concern about our results was that we might not have enough variability on the individual difference measures since all the subjects are Rice students, who are generally high ability students. For most of the individual difference measures and all the dependent measures, internal reliability was calculated as split-half (odd half/even half) correlation adjusted with the Spearman-Brown prophecy formula. For operation span, reading span, and verbal SAT, the internal reliability was obtained from previous studies (Friedman & Miyake, 2005, Unsworth, Heitz, Schrock & Engle, 2005). Table 8 shows the mean, standard deviation, range and reliability for the all the measurements. For the individual differences measures, while most of subjects performed well in all the tasks, their scores were distributed widely on each scale. The reliability of all these tests is very good. For the dependent measures, most variables had moderate to low reliability, with the lowest reliabilities for syntactic interference (reliability: 0.21) and semantic interference (reliability: 0.23) in the accuracy data and the semantic interference RT effect in the critical region (reliability: 0.28). Because of these low reliabilities for difference scores, it was important to evaluate the results

using a multiple regression approach as well.

Table 8. Reliability of all the measurements

Individual differences Measures	Index	Mean	Range	s.d.	Reliability
Operation span	Total score	63/75	36-75	8.5	0.78*** (Internal consistency)
Reading span	Total score	61/75	26-75	10.6	0.83-0.92*** (Split-half)
Verbal SAT	Total score	700	520-800	76	0.93*** (Internal consistency)
Category probe	Accuracy	0.81	0.67-0.94	0.10	0.74***
Digit span	Accuracy	0.74	0.43-1.00	0.10	0.73***
Stroop	RT (ms)	113	12-265	57.6	0.89***
Vocabulary	Score	36	19-44	5.4	0.82***
Dependent measures					
Syn/ critical	RT (ms)	-2.91	-449-368	117.778	0.40*
Sem/ critical	RT (ms)	14.6	-305-429	112.05	0.28
Syn/ spillover	RT (ms)	44.32	-321-380	117.784	0.34*
Sem/ spillover	RT (ms)	27.66	-361-491	119.477	0.57***
Syn/ question	RT (ms)	32.37	-462-499	144.96	0.51***
Sem/ question	RT (ms)	132.14	-143-605	148.954	0.51***
Syn/ accuracy	Accuracy	0.0396	-0.13-0.25	0.07969	0.21
Sem/ accuracy	Accuracy	0.1115	-0.3-0.5	0.15736	0.23
Overall accuracy	Accuracy	0.8628	0.6-0.99	0.07887	0.79***

*$p < .05$.
**$p < .01$.
***$p < .001$.

Note. Syn refers to syntactic interference effect size. Sem refers to semantic interference effect size. The reliability of operation span, reading span and verbal SAT was obtained from previous studies (Friedman & Miyake, 2005, Unsworth et al., 2005).

The correlations, which reached significance, between all the individual differences measures and interference effect sizes are shown in Table 9 and the scatterplots are shown in Figure 6. Appendix C provides the full correlation matrix.

Table 9. Correlations between interference effect size and individual difference measures

Model	Tasks	Interference (effect size)	Regions	*r*	Control variable	
					Vocabulary	SAT
Multiple capacities	Digit span	None	All	*p*s > .05	*p*s > .05	*p*s > .05
	Category probe	Sem/RT	Critical Question	-.24* -0.27**	-.24* -.23*	-.26** -.23*
General vs. Sentence-specific WMC	Reading span	Sem/Accuracy		-.24*	-.30**	-.26*
		Syn/RT	Spillover	-.23*	-.25**	-.17 (.11)
	Operation span	Sem/ Accuracy		-.21*	-.24*	-.22 *
Retrieval interference	Stroop	Sem/ RT	Question	.20 (.05)	.17 (.10)	.14 (.18)
	Vocabulary	Sem/RT	Question	-.26**		-.12 (.24)

* $p < .05$.
** $p < .01$.

Note. Values in parentheses are nonsignificant *p* values. "RT" refers to interference effect size in reaction time, and "Accuracy" refers to interference effect size in accuracy data. "Sem" refers to semantic interference effect size, and "Syn" refers to syntactic interference effect size.

In order to better compare the correlational results to predictions from different accounts, the results are organized by models in Table 9. First, the multiple capacities account predicted correlations between semantic interference effects and semantic STM, but no correlations between phonological STM and interference effects. Our results confirmed such predictions: accuracy on the category probe task was negatively related to the degree of semantic interference in self-paced reading times in the critical region (r = -.24, p = .02), and the degree of semantic interference effect in the speed of question answering (r = -.27, p = .001). However, digit span

failed to correlate with any of the interference effects (*ps* > .37). In addition, neither the semantic STM measures nor the phonological STM measure correlated with any of the syntactic interference effects.

Second, the general WMC account predicted correlations between complex span tasks and all of the interference effects. Some correlations with reading span and operation span were obtained: 1) reading span was negatively related to syntactic interference for RTS in the spillover region (r = -.23, p = .02) and semantic interference in comprehension question accuracy (r = -.24, p = .02);, 2) operation span was negatively correlated with semantic interference in comprehension question accuracy as well (r = -.21, p = .04). However, neither complex span measure was related to semantic interference in the self-paced reading measure or to syntactic interference in the comprehension question measures. Reading span but not operation span was related to syntactic interference in self-paced reading. Thus, support for the general WMC approach was only partial.

The correlations between the semantic STM measure and semantic interference in the critical region and between reading span and syntactic interference in the spillover region are contradictory to the predictions from sentence-specific WMC account (Caplan & Waters, 1999). That is, this account only predicted correlations with span measures in the off-line question answering measures but not in the online measures.

Third, the retrieval based account hypothesized correlations between interference effect size and Stroop and vocabulary, which again would be expected for all measures. Our results showed that only the semantic interference effect size in the speed of question answering was correlated with vocabulary (r = -.26, p = .001), and with the Stroop effect (r = .20, p = .05).

As mentioned in section 2.2.2. about why we included a vocabulary test as well as the verbal SAT in the current experiment, some researchers have suggested that the quality of language representations is the source of individual differences in sentence processing ability rather than WM capacity (MacDonald & Christiansen, 2002; Perfetti, 2007; Van Dyke, Clinton & Anuenue, 2010). In order to control for possible differences in semantic knowledge that might contribute to the correlations reported above, we partialled out vocabulary, with the results shown in the second column from the right in Table 9. We also partialled out verbal SAT scores as a proxy for IQ; however, as noted earlier, there are high correlations between WM and IQ measures, and thus the interpretation of these partial correlations is problematic. Three participants had SAT rather than ACT scores and these ACT scores were converted to SAT scores by ACT-SAT concordance provided by the ACT website (http://www.act.org/aap/concordance/). The results are shown in the rightmost column of Table 9.

The results revealed that after controlling for vocabulary, all the correlations remained significant except the correlation between Stroop and semantic interference in speed of question answering RT ($r = .17, p = .10$). After controlling for verbal SAT score, most of the correlation tests remained significant except for the correlation between reading span and syntactic interference in self-paced reading time in the spillover region and ($r = -.17, p = .11$), and the correlation between semantic interference in speed of question answering with Stroop ($r = .14, p = .18$) and vocabulary ($r = -.12, p = .24$).

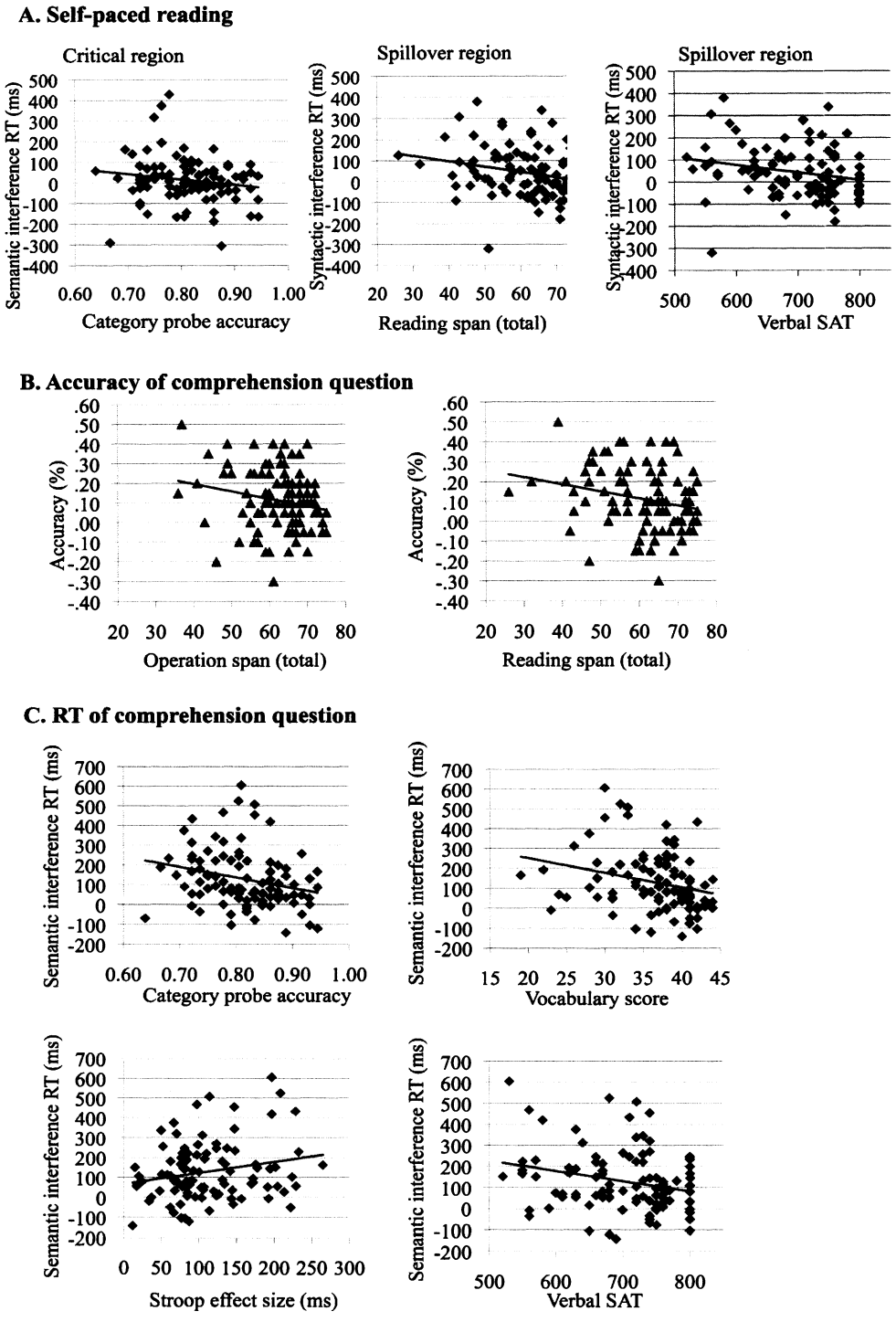


Figure 6. Correlations between interference effect size and other measurements.
All the correlations displayed here are significant. Figure 6(A) shows the significant correlation between interference effect size in self-paced reading time with reading span and verbal SAT. Figure 6(B) shows the correlation between interference effect size in accuracy of comprehension question with reading span and operation span. Figure 6(C) shows the correlation between interference effect size in speed of question answering and individual measures.

2.3.4 Multiple regression approach

There have been debates about the appropriateness of difference scores (or "raw change", "raw gain") vs. residuals in the analysis of covariance in the study of individual differences (Cronbach & Furby's, 1970). The analyses reported above implemented a difference score approach. However, some researchers (Cronbach & Furby, 1970; Lord, 1956) have pointed out that that "difference scores tend to be much more unreliable than the scores themselves", no matter how they may be adjusted. Cronbach and Furby (1970) claimed that it is more straightforward to ask about the regression of performance in condition Y (e.g., the high interference condition) on performance in condition X (e.g., the low interference condition) and other predictors of interest (e.g., semantic STM capacity). The significance of the regression weights for the predictors of interest indicates how well these measures predict performance beyond that accounted for by the baseline condition. Although difference scores can be reliable under certain conditions (Zimmerman & Wiliams, 1982; Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1983' Rogosa, Brandt, and Zimowski, 1982) the reliabilities of many of the difference score measures in the current study were not very high. Thus, we ran multiple regressions on the data here to determine whether these two analyses would provide similar results, by using performance in the more difficult condition (e.g., the high syntactic interference condition) as the dependent measure and performance in the easier condition (e.g., the low syntactic interference condition) as a predictor with the other variables of interest (e.g., reading span) as other predictors. For the Stroop effect, we added RTs in the neutral condition (asterisk) and RTs in the incongruent condition as independent measures separately in the equations. The results are shown in Table 10. Most of the results remained the same as what was obtained from differences score measures,

except: 1) the relation between reading span and syntactic interference effect size in the spillover region (*B*-value = -22.29, p = .07) became marginally significant; 2) for the comprehension questions, the mean RT in the incongruent condition in Stroop task was marginally related to the speed of question answering (*B*-value = .48, p = .08). The two rightmost columns show the significance of the B-weights when controlling for vocabulary and VSAT. After controlling for vocabulary, the correlation between reading span and syntactic interference effect size in the spillover region became significant (*B*-value = -26.60, p = .03). The changes in significance when controlling for VSAT were the same as those observed in the difference score analysis.

Table 10. Multiple regressions on the interference effect and individual difference measures

Model	Tasks	Interference (effect size)	Regions	*B*-value	Control variable	
					Vocabulary (*B*-value)	SAT (*B*-value)
Multiple capacities	Digit span	None	All	*p*s > .05	*p*s > .05	*p*s > .05
	Category probe	Sem/RT	Critical Question	-25.05* -38.15**	-24.77* -33.38*	-28.63* -33.46*
General vs. Sentence-specific WMC	Reading span	Sem/Accuracy		0.03**	0.03***	0.03**
		Syn/RT	Spillover	-22.29 (.07)	-26.60*	-12.50 (.21)
	Operation span	Sem/ Accuracy		0.02**	0.02**	0.02*
Retrieval interference	Stroop	Sem/ RT	Question	Neutral: -0.43 (.25) Incongruent: 0.48 (.08)	Neutral: -0.25 (.51) Incongruent: 0.39 (.15)	Neutral: -0.22 (.56) Incongruent: 0.36 (.19)
	Vocabulary	Sem/RT	Question	-33.35*		-21.48 (.26)

*p < .05.
**p < .01.
***p < .001.

Note. Values in parentheses are nonsignificant p values. "RT" refers to interference effect size in reaction time, and "Accuracy" refers to interference effect size in accuracy data. "Sem" refers to semantic interference effect size, and "Syn" refers to syntactic interference effect size.

2.3.5 Logarithmic transformation

One of the challenges to working memory capacity accounts has been the claim that processing speed accounts for the correlation between working memory and other capacities, such as general fluid intelligence (Fry & Hale, 1996). Some researchers have argued that processing speed is a general characteristic that influences other abilities. According to this account, high-span subjects are simply faster processors than low-span subjects in all tasks, including reading comprehension tasks (Kail & Salthouse, 1994; Salthouse, 1996). In general, larger reaction time effects are observed for slower subjects (Verhaeghen & Meersman, 1998). WM capacity could be reduced because slower speed results in less access to semantic information about words during list encoding and fewer rehearsals of words to keep them activated. Thus, correlations between sentence processing interference effects and WM or Stroop effects might simply be due to the fact that slower subjects generally show larger effects and smaller WM capacity. Although some prior studies have indicated that correlations between WM and other measure cannot be attributed to processing speed (Conway, Bunting, Therriault, and Minkoff, 2002) it was important to address whether processing speed contributed to the effects reported here.

In the current study, in order to take general slowing into account, a logarithmic transformation (Kirk, 1968) was performed on the RT data (Verhaeghen & Meersman, 1998). The log transformed RT were analyzed via a 2 (high or low semantic interference) × 2 (high or low syntactic interference) factorial repeated measures ANOVA as we did before. The ANOVA demonstrated a main effect of syntactic interference effect (F (1, 95) = 10.18, p = .002) and a main effect of semantic interference effect (F (1, 95) = 6.32, p = .01) in the spillover region. There was also a main effect of syntactic interference effect (F (1, 95) = 6.34, p = .01) and a

main effect of semantic interference effect (F (1, 95) = 82.85, p < .001) in the speed of question answering. The interactions were not significant. These results are the same as those obtained from the analysis on raw RT data. The results of correlational analyses are shown in Table 11. Most of the correlations survived log-transformation, except for the correlation between Stroop and speed of comprehension question answering (r = -.24, p = .20). The last two columns in Table 11 showed the partial correlation results after controlling for vocabulary or verbal SAT. The significance of the results remained the same after controlling for vocabulary. However, after controlling for verbal SAT, several correlations become non-significant. There was no significant correlation between syntactic interference effect size in spillover region and reading span score (r = -.17, p = .10). In addition, there was no significant correlation between semantic interference effect size in speed of question answering with Stroop (r = .08, p = .45) or vocabulary (r = -.09, p = .40).

Table 11. Correlations between interference effect size and individual differences on log-transformed data

Model	Tasks	Interference (effect size)	Regions	r value	Control variable	
					Vocabulary	SAT
Multiple capacities	Digit span	None	All	ps > .05	ps > .05	ps > .05
	Category probe	Sem/RT	Critical Question	-0.26** -0.24*	-0.25* -0.25*	-0.28** -0.25*
General vs. Sentence-specific WMC	Reading span	Syn/RT	Spillover	-0.22*	-0.22*	-0.17 (.10)
Retrieval interference	Stroop	Sem/ RT	Question	0.13 (.20)	0.10 (0.33)	0.08 (.45)
	Vocabulary	Sem/RT	Question	-0.22*		-0.09 (.40)

* p < .05. ** p < .01. *** p < .001.

Note. Values in parentheses are nonsignificant p values. "RT" refers to interference effect size in reaction time, and "Accuracy" refers to interference effect size in accuracy data. "Sem" refers to semantic interference effect size, and "Syn" refers to syntactic interference effect size.

2.3.6 Processing measures

In addition to the claim that semantic knowledge or general processing speed could serve as source of individual differences in sentence comprehension, there have also been debates about whether the task-specific processing skills required in the complex span tasks are the real cause of individual differences in performance on these tasks (Daneman & Tardiff, 1987, cited by Engle, Cantor & Carullo, 1992). However, it appears the task-specific processing portions of the WM tasks (e.g. reading skill in reading span task) are not important to the ability of WM tasks to predict performance on higher-order tasks (Cantor, Engle, & Hamilton, 1991; Engle, Nations, & Cantor, 1990; Engle, Cantor, & Carullo, 1992; Turner & Engle, 1986, 1989). However, none of these prior studies has specifically looked at syntactic and semantic interference effects as predicted by WM capacity. Thus, in the current experiment, we examined whether working memory capacity measures contribute independently after controlling for task-specific processing ability. As Engle et al. (1992) suggested in their experiment, we suggested that the sentence processing measures (that is, sentence reading time and sentence judgment time) from the reading span task could be interpreted as reflecting, at least to some extent, basic syntactic knowledge. In order to provide a comparison with another processing measure that is independent of syntactic processing, we also obtained similar measure of arithmetic ability (equation reading time and judgment time) from the operation span task. Thus, we examined whether operation span contributed independently to any of the sentence comprehension measures after controlling for arithmetic ability as reflected in these measures.

For each subject, all the correctly answered trials from the 75 observations in the reading and operation span tasks contributed to these mean viewing time scores

and mean judgment time scores, resulting in two mean scores for each subject after collapsing over group size of the trials. For all the subjects, the mean viewing time for operation span was 1991 ms (*S.D.* = 666 ms) and the mean judgment time was 984 ms (*S.D.* = 177 ms). The mean viewing time for reading span was 2710 ms (*S.D.* = 773 ms) and the mean judgment time was 752 ms (*S.D.* = 173 ms). If processing efficiency is an important variable in distinguishing high- from low- span subjects in sentence comprehension task, then the previously observed correlations between complex span tasks and interference effects should disappear after controlling for processing efficiency. The correlational results are shown in Table 12.

Table 12. Correlations between interference effect size and individual difference measures after controlling reading skill or arithmetic skill

	Region	Effect size	Control Variables	
			Reading skill	Arithmetic skill
Self-Paced Reading RTs	Spillover	Syntactic	Reading span, $r_1 = -.23, p = .03$, $r_2 = -.21, p = .04$	
Comprehension question accuracy	Question	Semantic	Reading span, $r_1 = -.25, p = .02$, $r_2 = -.24, p = .02$	Operation span, $r_1 = -.21, p = .044$, $r_2 = -.21, p = .04$

Note. The r_1 values are the results of analyses after controlling viewing time, and the r_2 values are the results of analyses after controlling judgment time.

After partialling out processing time, the correlations between complex span and interference effects were still significant. These results suggest that processing efficiency is not important to the relationship found between complex spans and interference resolution, and imply that the relationship between working memory and sentence comprehension exists over and above the relationship that span and sentence processing shared with processing efficiency.

2.3.7 Trade-off effects

We are also interested in whether there were tradeoffs between the semantic or syntactic interference effects across different regions. It is possible that different subject may apply different strategies during the experiment, e.g. some may have tried to resolve the semantic interference immediately when they processed the critical verb whereas others may have waited until the spillover region. If so, then a small interference effect in one region may result in a large interference effect in the subsequent region, or vice versa. To examine such possible tradeoffs, correlations were computed between the interference effect size in different regions on RTs, log-transformed RTs, and accuracy. The significant results are shown in Table 13. There were several positive correlations which indicated that the subjects showed larger semantic or syntactic interference effects in one region tended to show larger semantic or syntactic interference effect in other regions. Overall, we suggested that there were no remarkable tradeoff in the current experiment.

Table 13. Correlations between interference effect sizes across different regions separately for syntactic and semantic interference

Regions	Syntactic interference				Semantic interference	
	Critical	Spillover	Question	Critical	Spillover	Question
Critical		$r_1 = .22, p = .03$ $r_2 = .27, p = .001$			$r_1 = .27, p = .01$ $r_2 = .18, p = .08$	
Spillover						$r_1 = .24, p = .02$ $r_2 = .26, p = .01$
Question						

Note. Syn and Sem refer to syntactic interference effect size and semantic interference effect size. The r_1 values are the results of analyses on raw RT, and the r_2 values are the results of analyses on log-transformed data.

Summary. In section 2.3.2 to section 2.3.6, I tried several different analyses of the RT data. All these analyses resulted in highly similar results. The correlations of most interest, e.g. the correlations of online sentence comprehension measures with

category probe (for semantic interference) and with reading span (for syntactic interference) remained significant in all of the analyses, after partialling out vocabulary. Some of the correlations became nonsignificant after controlling for verbal SAT/vocabulary, or log-transforming the RT data. However, the observed changes were not particularly relevant to our research questions. For example, the non-significant correlation between interference effect size and the reading span task after controlling for verbal SAT could be explained on the grounds that these two measures are correlated (in the current experiment, $r = .35$, $p < .001$; in Daneman & Carpenter, 1983, $r = 0.46$, $p < .05$,). As for the reduction of the correlation between Stroop and interference effect size after controlling for verbal SAT or vocabulary, one possible explanation is that there is an association between the quality of a lexical representations and interference resolution (MacDonald & Christiansen, 2002, Van Dyke, Johns, & Kunona, 2010). However, the support for a role of the quality of lexical represntations in the semantic and syntactic interference effects was very weak in the current experiment. This will be discussed further in the discussion section. Given the similarity of the results for different ways of running the analyses, we feel that we can place considerable confidence in the major results of interest as shown in section 2.3.2.

3 General discussion

Motivated by the controversial issue about how to explain the nature of the capacity limit during sentence comprehension, our current project looked into this question by relating interference resolution ability during sentence processing to performance on a set of individual differences measures. Previous studies postulated a correlation between sentence comprehension ability and general verbal WM (Fedorenko et al., 2006, 2007; Just & Carpenter, 1992; Gordon et al., 2001, 2002; Kane et al., 2001), semantic STM (Martin et al., 1994, 2004), or knowledge differences (MacDonald & Christiansen, 2002; Van Dyke et al., 2010, Perfetti, 2007). All these predictions from different accounts have been confirmed in offline sentence processing measures. However, for online sentence comprehension, the current data provided strong evidence for the multiple capacities account that only semantic STM, but neither phonological STM, general WM, vocabulary, nor interference resolution ability could predict participants' performance. The impact of our finding on understanding individual differences in sentence comprehension is that future studies focus on this topic should look into the separable retention capacities of STM.

Sentence processing

With the revised materials, we successfully eliminated the local coherence effect found in the pilot study and replicated both syntactic and semantic interference effects reported in earlier studies (Van Dyke, 2007). However, the time course of the syntactic effect was delayed when compared to the results from Van Dyke's (2007) eye-tracking experiment, in which syntactic interference was obtained in the critical region. In the phrase-by-phrase reading paradigm, effects sometimes show up downstream relative to where they are predicted, most likely because subjects get ahead of themselves in button pressing to advance the presentation (Just, Carpenter, &

Wooley, 1982). This could explain why the semantic and syntactic interference effects did not show up in the critical region but in the spillover region. Unlike the case in Van Dyke's (2007) study, a time course difference between semantic and syntactic interference effects was not very obvious in our experiment. Although the syntactic interference effect size was larger than the semantic interference effect in the spillover region and the semantic interference effect size was larger than the syntactic interference effect size in the speed of question answering, both interference effects were significant in the two regions. As mentioned earlier, this could be explained by our greater power, given that we had almost triple the sample size of Van Dyke's experiment.

Relations between sentence processing and individual difference measures

To examine the relationship between interference effect size and individual differences measurements, we conducted multiple different analyses, including correlational analyses, residual analyses, analyses on log-transformed data, and partial correlational analyses by controlling vocabulary, SAT or processing efficiency. All these analyses resulted in highly similar results. Based on these findings, we have reached several conclusions. First, our results were most consistent with the multiple capacities account proposed by Martin and colleagues (Martin & He, 2004) in which only certain components of working memory capacity underlie sentence comprehension; specifically, semantic short-term memory but not phonological short-term memory is critical for sentence processing. In the current experiment, digit span performance, which was assumed to primarily reflect phonological retention, failed to correlate with any of the interference measures (all $ps > .30$), while category probe accuracy correlated with the semantic interference effect size in the critical region and in the question region. Thus, individuals with better semantic short-term memory

performed better in semantic interference resolution, presumably due to better semantic retention ability. Also, the fact that syntactic interference effect was not correlated with either semantic STM or phonological STM, supported Martin and Romani's finding (1994) that phonological, semantic, and syntactic retention abilities are independent components, suggesting a separate capacity for maintaining syntactic information.

Second, for the debates between general vs. specific online sentence-processing WM resources, our results revealed that there were correlations between complex span measures and interference effect size. Both reading span and operation span were related to semantic interference effect size in the accuracy data, and reading span was also related to syntactic interference effect size in the spillover region. Furthermore, when the variance common to the working memory capacity and processing efficiency was removed from the WM-interference effect relationship, WM capacity still served as a significant predictor of sentence comprehension. Although the correlations in question answering could possibly be attributed to post-interpretive processing in the Caplan and Waters (1999) framework, the correlation of syntactic interference with reading span in the spillover region is less easily accommodated in this fashion. Moreover, the correlation in the critical region of semantic interference with category probe accuracy also indicates a relation between online (interpretive) sentence processing and a standard span measure. Thus, the results are not compatible with Caplan and Waters' specific capacity framework. However, some of the results could be considered as consistent with the claims of Gibson et al. and Just and Carpenter that general verbal working memory capacities are involved in online sentence processing. Finally, we found some support for the two hypotheses from retrieval-based account with both vocabulary and resistance to

interference (as tapped by Stroop) correlated with interference effect size (though the correlation with the Stroop effect was only of marginal significance when factoring out vocabulary).

The question remains, however, as to whether any one of the approaches can provide a comprehensive account of all of the findings. If attention is restricted to the online sentence processing results, then the multiple capacities approach (Martin et al., 1994) provides the best account of the findings. The two significant correlations obtained there were a correlation between semantic interference and category probe in the critical region and a correlation between syntactic interference and reading span in the spillover region. It should be emphasized that the semantic interference effect correlated with the category probe measure and NOT with reading span, and the syntactic interference effect correlated with reading span and NOT with category probe. Thus, the size of the semantic interference effect appears to relate to a specific capacity for retaining semantic information. The correlation between syntactic interference and reading span might be explained on the grounds that the reading span task involves syntactic processing and thus could tap a specialized capacity for maintaining syntactic information. This hypothesis was partially supported by the result that the correlation between reading span and syntactic interference effect size remained significant after factoring out sentence processing efficiency (as measured by sentence processing time). As argued by Martin and Romani (1994), their patient data indicated a separation between the maintenance of syntactic and semantic information. Thus, the multiple capacities approach would be expanded to include a buffer for syntactic information, together with the assumption that both semantic and syntactic retention are important for sentence processing, whereas phonological retention is not. The correlations involving operation span and the Stroop effect in

question answering might be attributed to general reasoning or memory retrieval effects involved in post-interpretive processing, similar to the claims made by Caplan and Waters for offline sentence processing. All of the correlations involving vocabulary could be accommodated on the grounds that the quality of semantic representations affects the capacity for semantic retention (Martin et al., 1999). The general WMC approach has difficulty accounting for the specificity of the correlations between span measures and the semantic and syntactic interference effects.

With respect to cue-based retrieval, there was only weak support for the approach, which consisted of correlations of vocabulary and the Stroop effect with semantic interference in question answering. However, there is compelling evidence supporting the cue-based retrieval approach in sentence processing, including the interference effects obtained in our experiment and others (Lewis, Vasishth, & Van Dyke, 2003, Van Dyke & McElree, 2006, Van Dyke, 2007). The semantic interference effect is particularly striking as it appears even when the interfering material was syntactically unsuitable. Other findings also provide strong support for the approach: 1) difficulty in unambiguous sentence processing is caused by interference effects rather than the distance between sentence elements that have to be integrated (Van Dyke & Lewis, 2003), 2) an extraneous load affects comprehension when the load items have features shared with retrieval targets (Gordon et al., 2002; Fedorenko et al., 2006; Van Dyke & McElree, 2006). Based on this account, another possible approach to explain the dissociation of semantic and syntactic interference and their correlations with different individual differences measures would be to assume that information outside focus degrades differentially for semantic and syntactic information. The relative rapid degrading of semantic information as tapped by

category probe task affects semantic interference resolution. The lack of distance effects seems particularly problematic for capacity-based approaches (but see Bartek, Lewis for some evidence of distance effects in sentence processing). That is, one might expect that greater distance would result in greater decay of representations and influence the size of interference effects.

According to cue-based retrieval model, only a limited amount of information can be actively maintained in the focus of attention during sentence comprehension and information outside this focus must be retrieved. Thus, one would predict that individual differences in the ability to resolve interference in retrieval of information outside the focus would constrain language comprehension. There have been few studies which have examined individual differences in cue-based retrieval in either sentence processing or in general verbal working memory tasks. One study that took a cue-based retrieval approach to list recall was carried out by Daily, Lovett and Reader (2001). They proposed an ACT-R model to explain individual differences in list recall and concluded that individual differences could be accounted for as different amounts of attentional activation. In Daily et al.'s model, successful retrieval relied on the total amount of activation of the target chunk of information in memory, which is the sum of base-level activation and source activation (attentional activation). Base-level activation reflects the recency and frequency of prior access to this chunk, and source activation flows from the current goal (e.g., retrieve the item in the first serial position) to related nodes. After assuming equality of base level activation across subjects, Daily et al. attempted to fit individual subjects' data by varying either the total amount of source activation, decay rate, or the retrieval threshold. Their results showed that variation in source activation could best account for individual differences in recall level and serial position effects. Therefore, Daily suggested that it

is the differential ability to activate goal-relevant information that caused individual differences in list recall. People with a higher level of source activation performed better.

The Daily, Lovett, and Reder (2001) approach could also apply to language processing. Within this approach, base-level activation could be reflected by quality of language representations which would differ across subjects and which could be assessed by measures such as receptive vocabulary (MacDonald & Christiansen, 2002; Van Dyke et al., 2010, Perfetti, 2007) or syntactic knowledge. Some people may have more precise representations of words than others and may have more or less experience in processing complex sentence structures. Source activation could also vary across individuals in terms of their ability to activate goal-relevant information. The higher the source activation, the less susceptible the individual to interference. Based on the distinction between base-level activation and source activation, a possible explanation for the correlations we found in the experiment could be that all span measures reflect effects of base level activation and source activation. The category probe task required knowledge of word meanings and participants have to activate the item in the list that is in the same category in order to make the judgment. Both reading span and operation span task reflect basic reading or math ability as well as attentional control (Engle, 2002). In the current experiment, we included Stroop as a measure of resistance to retrieval interference. However, this may not be the most appropriate measure since we are looking at interference in memory and not interference from a predominant response (as in Stroop). There is evidence that proactive interference in memory involves a different type of interference resolution than resistance to a predominant response (Friedman & Miyake, 2004). Thus, a task such the recent negatives task developed by Monsell (1978), which involves proactive

interference (Rosen & Engle, 1998), might be a better method for assessing the relevant ability to activate goal-relevant information. In the recent negatives task, subjects need to respond to a probe word according to whether it appeared in the current list. A recent-negative probe trial is one in which the negative probe did not appear in the current list, but appeared in the list presented immediately preceding the current list. If source activation is a unitary ability that remains constant for an individual across tasks, we should be able to estimate source activation from span tasks (after accounting for knowledge differences underlying base activation differences) and then determine whether this measure of source activation accounts for individual differences in the susceptibility to interference in sentence processing. Actually, this assumption has been partially demonstrated by the finding that after controlling for reading proficiency, the remaining components of reading span, which was a more pure measure for attentional control ability, could still predict the individual differences in interference resolution. However, the sentence reading time within reading span task may not be the most reliable task to tap participants' basic knowledge of grammatical knowledge. In sum, according to retrieval-based approach, the pattern of individual differences in sentence processing measures should be accounted for by differences in source activation and differences in semantic and syntactic knowledge and by the relevant type of information involved in a particular effect (i.e., semantic knowledge but not syntactic knowledge is involved in semantic interference effects). In order for this approach to account for the selectivity we observed for semantic and syntactic interference and their correlations with different span measures, it would have to be the case that subject differences in syntactic and semantic knowledge led to these correlations. The fact that our correlations persisted after controlling for vocabulary does not lend much support to this notion. However,

perhaps testing individuals with a wider range of vocabulary abilities would reveal different results. Alternatively, it could be the case that the rate of decay of syntactic and semantic information differs across subjects, but if decay is involved, then effects of distance should be observed in sentence processing. Given that the Van Dyke and Lewis study is the only study directly examining distance vs. interference effects, further replication of their finding is needed.

4 Future Directions

As a future direction for this research (which is not part of this master's project), we would like to pursue the approach of modeling baseline and source activation separately. For example, for base-level activation, we could measure vocabulary to test base-level activation of semantic knowledge, and measure grammatical knowledge to test base-level activation of syntactic knowledge. For source activation, as mentioned in the discussion session, the recent negatives task might be a more appropriate task. Source level activation could be obtained from modeling individual differences in a span task. We could then determine if these measures if the same base level and source activation parameters that predict span measures could be used to predict sentence processing measures. If so, then the correlations between the sentence processing and span measures might be due to the commonality of these basic cognitive components across tasks.

References

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory (Handbook of perception and cognition) (2nd ed., pp. 237–313)*. San Diego, CA: Academic Press.

Caplan, D., & Waters, G.S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22,* 77–126.

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprhension. *Joournal of Memory and Language, 52*: 144-169.

Chen, E., West, W.C., Waters, G., & Caplan, D. (2006). Determinants of bold signal correlates of processing object-extracted relative clauses. *Cortex, 42,* 591–604.

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprhension. *Joournal of Memory and Language, 52*: 144-169.

Conway, A.R.A., Conwan, N., Bunting, M.F., Therriault, D.J., & Minkoff, S.R.B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163 – 183.

Cronbach, L.J., & Furby, L. (1970). How we should measure "change" or should we? *Psychological Bulletin, 74 (1)*, 68 – 80.

Daily, L.Z., Lovett, M.C., Reder, L.M. (2001). Modeling individual differences in working memory performance: a source activation account. *Cognitive Science, 25*: 315 – 353.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 561-583.

Daneman, M., & Carpenter, P. A. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18 (5),* 972-992.

Daneman. M., & Tardiff, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), Attention and performance XII (pp. 491-508). Hillsdale, NJ: Erlbaum.

Donlon, T. F. (1984). The College Board technical handbook for the SAT and Achievement Tests. New York: College Entrance Examination Board.

Evans, W., Waters, G., & Caplan, D. (2010). On-line interaction of mathematical and syntactic demand: failure to replicate. *CUNY conference.*

Engle, R.W., Cantor, J., & Carullo, J.J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of experimental psychology: learning, memory, and cognition, 18 (5)*, 972 – 992.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current directions in psychological science, 11 (1)*, 19 - 23.

Evan, W.S., Caplan, D., & Waters, G. (2010). Effects of concurrent arithmetical and syntactic complexity on self-paced reaction times and eye fixations. *Psychonomic bulletin & review*, (online first).

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of memory and language, 54 (4),* 541-553.

Fedorenko, E., Gibson, E., Rohde, D. (2007). The nature of working memory in linguistic, arithmetic and spatial integration processes. *Journal of memory and language, 56,* 246-269.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade. Psychological Science, 7, 237 – 241.

Gathercole, S.E., Short-term and working memory. (2001). Psychology Press.

Gibson E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition, 68*, 1-76.

Gordon, P.C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1411–1423.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13*, 425– 430.

Gordon, P.C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interfrence during language comprehension: evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(6),* 1304–1321.

Grodner, D.J. and Gibson, E.A. (2005). Consequences of the serial nature of linguistic input for sentential complexity. Cognitive Science. 29, 261–290

Hambrick, D.Z., & Engle, R.W. (2002). Effects of Domain Knowledge, Working Memory Capacity, and Age on Cognitive Performance: An Investigation of the Knowledge-Is-Power Hypothesis. *Cognitive Psychology, 4(44),* 339-287.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processing in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228–238.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*, 199 – 225.

Konieczny, L. (2000). Locality and parsing complexity. *Jorunal of psycholinguistic research, 29*, 627-645.

Lewis, R.L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research, 25*, 93-1 15.

Lewis, R.L. (1999). Accounting for the fine structure of syntactic working memory: Similarity-based interference as a unifying principle. Behavioral and Brain Sciences, 22, 105-106.

Lewis, R.L., Vasishth, S., & Van Dyke, J.A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10,* 447–454.

Lord, F.M. (1956). The measurement of growth. *Educational and psychological measurement, 16*, 421 – 437.

MacDonald, M.C., & Christiansen, M.H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological review, 109 (1)*, 35 – 54.

Martin, R.C., Romani. C. (1994a). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology, 8,* 506-523.

Martin, R. C., Shelton J. R. & Yaffee L. S. (1994b). Language Processing and Working Memory: Neuropsychological Evidence for Separate Phonological and Semantic Capacities. *Journal of memory and language*, *33(1),* 83-111.

Martin, R.C. (2003). Language processing: Functional organization and Neuroanatomical Basis. *Annual Review of Psychology, 54,* 55–89.

Martin, R.C., Wu, D., Jackson, D., Freedman, M. & Lesch, M. (2003). An event-related fMRI investigation of phonological vs. semantic short-term memory. *Journal of Neurolinguistics, 16*, 341–360.

Martin, R.C. (2005). Components of short-term memory and their relation to language processing. *Current direction in psychological science, 14 (4)*, 204-208.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General, 118,* 346–373.

McElree, B., & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General, 122,* 291–315.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48,* 67 – 91.

McElree, b. (2006). Accessing recent events. *The Psychology of Learning and Motivation, 46*, 155 – 200.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10*, 465-501.

Perfetti, C. (2007). Reading ability: lexical quality to comprehension. *Scientific studies of readings, 11(4)*, 357 – 383.

Rogosa, D. R., Brandt, D., & Zimowski, M. A. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 90*, 726 – 748.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of educational measurement, 20 (4)*, 335 – 343.

Rohde, T.E., & Thompson, L.A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83 – 92.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory & Language, 39*, 418-436.

Saffran, E. M., & Marin, O. S. M. (1975). Immediate memory for word lists and sentences in a patient with deficient auditory short-term memory. *Brain & Language, 2,* 420 – 433.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403 – 428.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language, 50,* 355–370.

Taylor, D., Lewis, R., & Badecker, W. (2008). Infividual difference in agreement planning: The role of working-memory control (poster).

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47,* 69–90.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R.W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.

Vallar, G., Baddeley, A. D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology, 1(2),* 121 – 141.

Van Dyke, J. A., Lewis, R. L. (2003) Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49,* 285 – 316.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 407–430.

Van Dyke, J.A., Johns, C.L., Kukona, A. (2010) Individual differences in sentence comprhension: a retrieval interference approach. CUNY conference 2010.

Verhaeghen, P. & De Meersman, L. (1998) Aging and the Stroop effects: a meta-analysis. *Psychology and aging, 13 (1)*, 120 – 126.

Zimmerman, D. W., & Williams, R. H. (1982) Gain score in research can be highly reliable. *Journal of educational measurement, 19*, 149 – 154.

Appendix A. Experimental sentences for pilot study

Item	Cond	Sentence	Question
1	a	The lady forgot that the client who had planned for the important meeting complained about the poor management at the firm.	Who complained about the management?
1	b	The lady forgot that the client who had planned for the important visitor complained about the poor management at the firm.	Who complained about the management?
1	c	The lady forgot that the client who declared that the meeting was important complained about the poor management at the firm.	Who complained about the management?
1	d	The lady forgot that the client who declared that the visitor was important complained about the poor management at the firm.	Who complained about the management?
2	a	The teacher noticed that the principal who had talked about the new curriculum responded to the recent budget cuts negatively.	Who responded to the budget cuts?
2	b	The teacher noticed that the principal who had talked about the new specialist responded to the recent budget cuts negatively.	Who responded to the budget cuts?
2	c	The teacher noticed that the principal who had mentioned that the curriculum is new responded to the recent budget cuts negatively.	Who responded to the budget cuts?
2	d	The teacher noticed that the principal who had mentioned that the specialist is new responded to the recent budget cuts negatively.	Who responded to the budget cuts?
3	a	The pilot remembered that the lady who was sitting in the smelly seat moaned about a refund for her ticket.	Who moaned about a refund?
3	b	The pilot remembered that the lady who was sitting near the smelly man moaned about a refund for her ticket.	Who moaned about a refund?
3	c	The pilot remembered that the lady who said that the seat was smelly moaned about a refund for her ticket.	Who moaned about a refund?
3	d	The pilot remembered that the lady who said that the man was smelly moaned about a refund for her ticket.	Who moaned about a refund?
4	a	The priest said that the woman who had stolen from the strict church departed from the building very quickly.	Who departed from the building?
4	b	The priest said that the woman who had stolen from the strict nun departed from the building very quickly.	Who departed from the building?
4	c	The priest said that the woman who discovered that the church is strict departed from the building very quickly.	Who departed from the building?
4	d	The priest said that the woman who discovered that the nun is strict departed from the building very quickly.	Who departed from the building?
5	a	The star discovered that the reporter who had waited for the black sedan lied about his position at the newspaper.	Who lied about his position?
5	b	The star discovered that the reporter who had waited for the black athlete lied about his position at the newspaper.	Who lied about his position?

5	c	The star discovered that the reporter who commented that the sedan was black lied about his position at the newspaper.	Who complained about the management?
5	d	The star discovered that the reporter who commented that the athlete was black lied about his position at the newspaper.	Who lied about his position?
6	a	The captain shouted that the soldier who had shot the hidden tank blundered in his aim terribly.	Who blundered in his aim?
6	b	The captain shouted that the soldier who had shot the hidden man blundered in his aim terribly.	Who blundered in his aim?
6	c	The captain shouted that the soldier who saw that the tank was hidden blundered in his aim terribly.	Who blundered in his aim?
6	d	The captain shouted that the soldier who saw that the man was hidden blundered in his aim terribly.	Who blundered in his aim?
7	a	The guard heard that the burglar who had given back the precious jewel whimpered about the consequence of his crime.	Who whimpered about the consequence?
7	b	The guard heard that the burglar who had given back the precious child whimpered about the consequence of his crime.	Who whimpered about the consequence?
7	c	The guard heard that the burglar who discovered that the jewel was precious whimpered about the consequence of his crime.	Who whimpered about the consequence?
7	d	The guard heard that the burglar who discovered that the child was precious whimpered about the consequence of his crime.	Who whimpered about the consequence?
8	a	The kid knew that the woman who had yelled about the dirty room whined about wanting to go out. frequently.	Who complained about the management?
8	b	The kid knew that the woman who had yelled about the dirty toddler whined about wanting to go out. frequently.	Who whined about wanting to go out?
8	c	The kid knew that the woman who yelled that the room was dirty whined about wanting to go out. frequently.	Who whined about wanting to go out?
8	d	The kid knew that the woman who yelled that the toddler was dirty whined about wanting to go out. frequently.	Who whined about wanting to go out?
9	a	The clerk saw that the teller who had worked in the boring room quit the job in a fit of anger.	Who quit the job?
9	b	The clerk saw that the teller who had worked for the boring boss quit the job in a fit of anger.	Who quit the job?
9	c	The clerk saw that the teller who complained that the room was boring quit the job in a fit of anger.	Who quit the job?
9	d	The clerk saw that the teller who complained that the boss was boring quit the job in a fit of anger.	Who quit the job?
10	a	The lady said that the boy who had dumped out the rich soil apologized for being so rude that day.	Who apologized for being rude?
10	b	The lady said that the boy who had dumped the rich girl apologized for	Who apologized for

		being so rude that day.	being rude?
10	c	The lady said that the boy who discovered that the soil was rich apologized for being so rude that day.	Who complained about the management?
10	d	The lady said that the boy who discovered that the girl was rich apologized for being so rude that day.	Who apologized for being rude?
11	a	The grocer remembered that the boy who had picked up the heavy delivery groaned about the bulky load in his arms.	Who groaned about the bulky load?
11	b	The grocer remembered that the boy who had picked up the heavy child groaned about the bulky load in his arms.	Who groaned about the bulky load?
11	c	The grocer remembered that the boy who commented that the delivery was heavy groaned about the bulky load in his arms.	Who groaned about the bulky load?
11	d	The grocer remembered that the boy who commented that the child was heavy groaned about the bulky load in his arms.	Who groaned about the bulky load?
12	a	The woman saw that the girl who had come with the cute teddy bear cried over the toy in her room.	Who cried over the toy?
12	b	The woman saw that the girl who had come with the cute baby cried over the toy in her room.	Who cried over the toy?
12	c	The woman saw that the girl who said that the teddy bear was cute cried over the toy in her room.	Who cried over the toy?
12	d	The woman saw that the girl who said that the baby was cute cried over the toy in her room.	Who cried over the toy?
13	a	The woman saw that the waitress who had cleaned up the messy table chatted with some old friends.	Who complained about the management?
13	b	The woman saw that the waitress who had cleaned near the messy guest chatted with some old friends.	Who chatted with friends?
13	c	The woman saw that the waitress who ignored that the table was messy chatted with some old friends.	Who chatted with friends?
13	d	The woman saw that the waitress who ignored that the guest was messy chatted with some old friends.	Who chatted with friends?
14	a	The magician saw that the lady who had believed in the mystical ball smirked at the simple trick in the corner.	Who smirked at the simple trick?
14	b	The magician saw that the lady who had believed in the mystical man smirked at the simple trick. in the corner.	Who smirked at the simple trick?
14	c	The magician saw that the lady who believed that the ball was mystical smirked at the simple trick. in the corner.	Who smirked at the simple trick?
14	d	The magician saw that the lady who believed that the man was mystical smirked at the simple trick. in the corner.	Who smirked at the simple trick?
15	a	The child thought that the mother who had yelled at the bad dog overreacted a bit to the mistake.	Who overreacted to the mistake?

15	b	The child thought that the mother who had yelled at the bad boy overreacted a bit to the mistake.	Who overreacted to the mistake?
15	c	The child thought that the mother who knew that the dog was bad overreacted a bit to the mistake.	Who responded to the budget cuts?
15	d	The child thought that the mother who knew that the boy was bad overreacted a bit to the mistake.	Who overreacted to the mistake?
16	a	The man asked if the student who had trained in the difficult program agreed with the new policy at the university.	Who might have agreed with the policy?
16	b	The man asked if the student who had trained with the difficult professor agreed with the new policy at the university.	Who might have agreed with the policy?
16	c	The man asked if the student who complained that the program was difficult agreed with the new policy at the university.	Who might have agreed with the policy?
16	d	The man asked if the student who complained that the professor was difficult agreed with the new policy at the university.	Who might have agreed with the policy?
17	a	The lady saw that the nanny who is devoted to the quiet home cried bitterly about the upcoming move.	Who cried bitterly?
17	b	The lady saw that the nanny who is devoted to the quiet child cried bitterly about the upcoming move.	Who cried bitterly?
17	c	The lady saw that the nanny who loves that the home is quiet cried bitterly about the upcoming move.	Who cried bitterly?
17	d	The lady saw that the nanny who loves that the child is quiet cried bitterly about the upcoming move.	Who cried bitterly?
18	a	The guide recalled that the tourist who had searched for the legendary city hiked through the foothills daily.	Who responded to the budget cuts?
18	b	The guide recalled that the tourist who had searched for the legendary Indian hiked through the foothills daily.	Who hiked through the foothills?
18	c	The guide recalled that the tourist who discovered that the city was legendary hiked through the foothills daily.	Who hiked through the foothills?
18	d	The guide recalled that the tourist who discovered that the Indian was legendary hiked through the foothills daily.	Who hiked through the foothills?
19	a	The senior wondered if the aide who had picked up the lost wallet arrived home at a decent hour.	Who arrived home?
19	b	The senior wondered if the aide who had picked up the lost child arrived home at a decent hour.	Who arrived home?
19	c	The senior wondered if the aide who saw that the wallet was lost arrived home at a decent hour.	Who arrived home?
19	d	The senior wondered if the aide who saw that the child was lost arrived home at a decent hour.	Who arrived home?
20	a	The man knew that the merchant who had sold the creative artwork	Who lied about the

		lied about the selling price of the item.	price?
20	b	The man knew that the merchant who had sold to the creative artist lied about the selling price of the item.	Who lied about the price?
20	c	The man knew that the merchant who said that the artwork was creative lied about the selling price of the item.	Who responded to the budget cuts?
20	d	The man knew that the merchant who said that the artist was creative lied about the selling price of the item.	Who lied about the price?
21	a	The neighbor saw that the mailman who was afraid of the angry dog knocked very softly at the door.	Who knocked at the door?
21	b	The neighbor saw that the mailman who was afraid of the angry woman knocked very softly at the door.	Who knocked at the door?
21	c	The neighbor saw that the mailman who wondered if the dog was angry knocked very softly at the door.	Who knocked at the door?
21	d	The neighbor saw that the mailman who wondered if the woman was angry knocked very softly at the door.	Who knocked at the door?
22	a	The driver saw that the peddler who had begged for the kind handout paid with a crisp ten dollar bill.	Who paid?
22	b	The driver saw that the peddler who had begged to the kind lady paid with a crisp ten dollar bill.	Who paid?
22	c	The driver saw that the peddler who mentioned that the handout was kind paid with a crisp ten dollar bill.	Who paid?
22	d	The driver saw that the peddler who mentioned that the lady was kind paid with a crisp ten dollar bill.	Who paid?
23	a	The teen said that the boy who had asked for the popular candy smiled sweetly during class.	Who responded to the budget cuts?
23	b	The teen said that the boy who had asked about the popular girl smiled sweetly during class.	Who smiled sweetly?
23	c	The teen said that the boy who implied that the candy was popular smiled sweetly during class.	Who smiled sweetly?
23	d	The teen said that the boy who implied that the girl was popular smiled sweetly during class.	Who smiled sweetly?
24	a	The collector heard that the owner who had talked about the interesting painting concurred with the museum's generous appraisal.	Who concurred with the appraisal?
24	b	The collector heard that the owner who had talked with the interesting painter concurred with the museum's generous appraisal.	Who concurred with the appraisal?
24	c	The collector heard that the owner who said that the painting was interesting concurred with the museum's generous appraisal.	Who concurred with the appraisal?
24	d	The collector heard that the owner who said that the painter was interesting concurred with the museum's generous appraisal.	Who concurred with the appraisal?

25	a	The boy noticed that the elf who had talked about the amazing cave laughed at the little white lie.	Who laughed at the white lie?
25	b	The boy noticed that the elf who had talked about the amazing genie laughed at the little white lie.	Who laughed at the white lie?
25	c	The boy noticed that the elf who swore that the cave was amazing laughed at the little white lie.	
25	d	The boy noticed that the elf who swore that the genie was amazing laughed at the little white lie.	Who laughed at the white lie?
26	a	The man said that the teller who had worked on the new account cheated for almost six months last year.	Who cheated?
26	b	The man said that the teller who had worked with the new manager cheated for almost six months last year.	Who cheated?
26	c	The man said that the teller who admitted that the account was new cheated for almost six months last year.	Who cheated?
26	d	The man said that the teller who admitted that the manager was new cheated for almost six months last year.	Who cheated?
27	a	The cop saw that the lady who had screamed about the dangerous fire watched from the window on her second floor.	Who watched from the window?
27	b	The cop saw that the lady who had screamed about the dangerous robber watched from the window on her second floor.	Who watched from the window?
27	c	The cop saw that the lady who screamed that the fire was dangerous watched from the window on her second floor.	Who watched from the window?
27	d	The cop saw that the lady who screamed that the robber was dangerous watched from the window on her second floor.	Who watched from the window?
28	a	The editor knew that the critic who had raved about the memorable play laughed through most of the second act.	Who laughed?
28	b	The editor knew that the critic who had raved about the memorable actress laughed through most of the second act.	Who laughed?
28	c	The editor knew that the critic who said that the play was memorable laughed through most of the second act.	Who laughed?
28	d	The editor knew that the critic who said that the actress was memorable laughed through most of the second act.	Who laughed?
29	a	The man thought that the customer who had ordered the plain casserole misbehaved in an inexcusable manner that evening.	Who misbehaved in an inexcusable manner?
29	b	The man thought that the customer who had ordered from the plain waitress misbehaved in an inexcusable manner that evening.	Who misbehaved in an inexcusable manner?
29	c	The man thought that the customer who complained that the casserole was plain misbehaved in an inexcusable manner that evening.	Who misbehaved in an inexcusable manner?
29	d	The man thought that the customer who complained that the waitress	Who misbehaved in an

		was plain misbehaved in an inexcusable manner that evening.	inexcusable manner?
30	a	The ranger noticed that the hiker who was hiking along the unclear trail stumbled over the bulging tree root in the forest.	Who stumbled over the tree root?
30	b	The ranger noticed that the hiker who was hiking behind the unclear leader stumbled over the bulging tree root in the forest.	Who stumbled over the tree root?
30	c	The ranger noticed that the hiker who realized that the trail was unclear stumbled over the bulging tree root in the forest.	Who stumbled over the tree root?
30	d	The ranger noticed that the hiker who realized that the leader was unclear stumbled over the bulging tree root in the forest.	Who stumbled over the tree root?
31	a	The agent reported that the ambassador who had learned about the sinister conspiracy arrived at the embassy with a guard.	Who arrived at the embassy?
31	b	The agent reported that the ambassador who had learned about the sinister terrorist arrived at the embassy with a guard.	Who arrived at the embassy?
31	c	The agent reported that the ambassador who recognized that the conspiracy was sinister arrived at the embassy with a guard.	Who arrived at the embassy?
31	d	The agent reported that the ambassador who recognized that the terrorist was sinister arrived at the embassy with a guard.	Who arrived at the embassy?
32	a	The commander wondered whether the soldiers who were searching for the captured supplies prevailed in the vital intelligence mission eventually.	Who prevailed?
32	b	The commander wondered whether the soldiers who were searching for the captured pilots prevailed in the vital intelligence mission eventually.	Who prevailed?
32	c	The commander wondered whether the soldiers who suspected that the supplies were captured prevailed in the vital intelligence mission eventually.	Who prevailed?
32	d	The commander wondered whether the soldiers who suspected that the pilots were captured prevailed in the vital intelligence mission eventually.	Who prevailed?
33	a	The band discovered that the manager who had worked out the clever show boozed at the bar during intermission.	Who boozed at the bar?
33	b	The band discovered that the manager who had worked with the clever producer boozed at the bar during intermission.	Who boozed at the bar?
33	c	The band discovered that the manager who said that the show was clever boozed at the bar during intermission.	Who boozed at the bar?
33	d	The band discovered that the manager who said that the producer was clever boozed at the bar during intermission.	Who boozed at the bar?
34	a	The politician found that the journalist who had written about the controversial product responded in the paper's editorial section.	Who responded in the editorial section?

34	b	The politician found that the journalist who had written about the controversial spokesman responded in the paper's editorial section.	Who responded in the editorial section?
34	c	The politician found that the journalist who wrote that the product was controversial responded in the paper's editorial section.	Who responded in the editorial section?
34	d	The politician found that the journalist who wrote that the spokesman was controversial responded in the paper's editorial section.	Who responded in the editorial section?
35	a	The architect regretted that the contractor who was concerned about the strong roof objected rudely with a loud voice.	Who objected rudely?
35	b	The architect regretted that the contractor who was concerned about the strong plumber objected rudely with a loud voice.	Who objected rudely?
35	c	The architect regretted that the contractor who doubted that the roof was strong objected rudely with a loud voice.	Who objected rudely?
35	d	The architect regretted that the contractor who doubted that the plumber was strong objected rudely with a loud voice.	Who objected rudely?
36	a	The judge realized that the attorney who had commented about the unusual motion relented after the cross-examination unwillingly.	Who relented after the cross-examination?
36	b	The judge realized that the attorney who had commented about the unusual witness relented after the cross-examination unwillingly.	Who relented after the cross-examination?
36	c	The judge realized that the attorney who commented that the motion was unusual relented after the cross-examination unwillingly.	Who relented after the cross-examination?
36	d	The judge realized that the attorney who commented that the witness was unusual relented after the cross-examination unwillingly.	Who relented after the cross-examination?
37	a	The social worker was surprised that the resident who said that the warehouse was dangerous complained about the investigation all night.	Who complained about the investigation?
37	b	The social worker was surprised that the resident who said that the neighbor was dangerous complained about the investigation all night.	Who complained about the investigation?
37	c	The social worker was surprised that the resident who was living in the dangerous warehouse complained about the investigation all night.	Who complained about the investigation?
37	d	The social worker was surprised that the resident who was living near the dangerous neighbor complained about the investigation all night.	Who complained about the investigation?
38	a	The major knew that the instructor who had looked for the prepared resume discovered the truth this morning.	Who discovered the truth?
38	b	The major knew that the instructor who had looked for the prepared student discovered the truth this morning.	Who discovered the truth?
38	c	The major knew that the instructor who assumed that the resume was prepared discovered the truth this morning.	Who discovered the truth?
38	d	The major knew that the instructor who assumed that the student was prepared discovered the truth this morning.	Who discovered the truth?
39	a	The manager saw that the handyman who thought that the project was	Who stood outside the

		ambitious stood outside the office for a long time.	office?
39	b	The manager saw that the handyman who thought that the foreman was ambitious stood outside the office for a long time.	Who stood outside the office?
39	c	The manager saw that the handyman who worked on the ambitious project stood outside the office for a long time.	Who stood outside the office?
39	d	The manager saw that the handyman who worked with the ambitious foreman stood outside the office for a long time.	Who stood outside the office?
40	a	The clerk noticed that the couple who had requested the sophisticated vase talked to the manager in the lobby.	Who talked to the manager?
40	b	The clerk noticed that the couple who had requested a sophisticated salesman talked to the manager in the lobby.	Who talked to the manager?
40	c	The clerk noticed that the couple who remarked that the vase was sophisticated talked to the manager in the lobby.	Who talked to the manager?
40	d	The clerk noticed that the couple who remarked that the salesman was sophisticated talked to the manager in the lobby.	Who talked to the manager?
41	a	The scientist read that the candidate who charged that the senator was dishonest lost the race unexpectedly.	Who lost the race?
41	b	The scientist read that the candidate who charged that the commercial was dishonest lost the race unexpectedly.	Who lost the race?
41	c	The scientist read that the candidate who was attacked by the dishonest senator lost the race unexpectedly.	Who lost the race?
41	d	The scientist read that the candidate who was attacked by the dishonest commercial lost the race unexpectedly.	Who lost the race?
42	a	The doctor knew that the specialist who thought that the disease is terminal discussed all the options in the meeting.	Who discussed all the options?
42	b	The doctor knew that the specialist who thought that the patient is terminal discussed all the options in the meeting.	Who discussed all the options?
42	c	The doctor knew that the specialist who has worked with the terminal disease discussed all the options in the meeting.	Who discussed all the options?
42	d	The doctor knew that the specialist who has worked with the terminal patient discussed all the options in the meeting.	Who discussed all the options?
43	a	The lawyer suggested that the judge who decided that the evidence was questionable misunderstood the facts for some reason.	Who misunderstood the facts?
43	b	The lawyer suggested that the judge who decided that the witness was questionable misunderstood the facts for some reason.	Who misunderstood the facts?
43	c	The lawyer suggested that the judge who had criticized the questionable evidence misunderstood the facts for some reason.	Who misunderstood the facts?
43	d	The lawyer suggested that the judge who had criticized the questionable witness misunderstood the facts for some reason.	Who misunderstood the facts?

44	a	The professor heard that the secretary who thought that the policy was unreasonable quit very angrily last month.	Who quit angrily?
44	b	The professor heard that the secretary who thought that the director is unreasonable quit very angrily last month.	Who quit angrily?
44	c	The professor heard that the secretary who was complaining about the unreasonable policy quit very angrily last month.	Who quit angrily?
44	d	The professor heard that the secretary who was complaining about the unreasonable director quit very angrily last month.	Who quit angrily?
45	a	The chiropractor swore that the girl who said that the pain was annoying visited the clinic yesterday.	Who visited the clinic?
45	b	The chiropractor swore that the girl who said that the mother was annoying visited the clinic yesterday.	Who visited the clinic?
45	c	The chiropractor swore that the girl who was complaining about the annoying pain visited the clinic yesterday.	Who visited the clinic?
45	d	The chiropractor swore that the girl who was complaining about the annoying mother visited the clinic yesterday.	Who visited the clinic?
46	a	The dancer heard that the director who claimed that the performance was outrageous wanted to quit immediately.	Who wanted to quit?
46	b	The dancer heard that the director who claimed that the performer was outrageous wanted to quit immediately.	Who wanted to quit?
46	c	The dancer heard that the director who disliked the outrageous performance wanted to quit immediately.	Who wanted to quit?
46	d	The dancer heard that the director who disliked the outrageous performer wanted to quit immediately.	Who wanted to quit?
47	a	The accountant thought that the developer who believed that the building was ruined spent too much money on this.	Who spent too much money?
47	b	The accountant thought that the developer who believed that the company was ruined spent too much money on this.	Who spent too much money?
47	c	The accountant thought that the developer who had bought the ruined building spent too much money on this.	Who spent too much money?
47	d	The accountant thought that the developer who had bought the ruined company spent too much money on this.	Who spent too much money?
48	a	The student remembered that the professor who declared that the lecture was intense recommended a graduate school to the audience.	Who recommended a graduate school?
48	b	The student remembered that the professor who declared that the lecturer was intense recommended a graduate school to the audience.	Who recommended a graduate school?
48	c	The student remembered that the professor had arrived after the lecture recommended a graduate school to the audience.	Who recommended a graduate school?
48	d	The student remembered that the professor had arrived after the	Who recommended a

		lecturer recommended a graduate school to the audience.	graduate school?
Fillers			
1	fl	The girl that the mad seamstress reported yesterday had searched for the exclusive fabric for many years.	Who searched for the fabric?
1	fl	The missing boy that the man accidentally found in the forest was sent back home by the police officer last week.	Who was sent back?
1	fl	The queen that the ridiculous jester easily amused realized that the singer was very popular at the banquet.	Who was very popular?
1	fl	The student that the judicial board acquitted had visited the loud party that day but left before the fight.	Who visited the loud party?
1	fl	The player that the energetic coach enthusiastically supported had played in the tough competition for more than ten years.	Who had played in the tough competition?
1	fl	The surgeon that young doctor routinely tested has operated on many difficult patients in the past.	Who operated on many difficult patients?
1	fl	The enemy that the wounded soldier quickly noticed had shot at the hidden man.	Who shot at the hidden man?
1	fl	The physicist that the senior scientist admired made too much noise during the lecture and was asked to keep quiet.	Who made noise?
1	fl	The teacher that the little girl trusted had unfairly graded the difficult test this time.	Who graded unfairly?
1	fl	The woman that the disorganized man openly criticized had figured out that the technician was unreliable.	Who figured out the technician was unreliable?
1	fl	The waitress that the truck driver saw knew that the menu for tonight was very greasy.	Who knew the menu was greasy?
1	fl	The naughty boy that the guidance counselor happily advised and invited for coffee dumped the rich girl.	Who dumped the girl?
1	fl	The customer that the busy grocer checked out cried that the bag was heavy.	
1	fl	The child that the teacher aggressively ignored complained that the tape was loud and unclear during the class.	
1	fl	The mother that the angry daughter rudely reproached paid for the filthy dress without hesitation.	
1	fl	The man that the fraternity brother criticized thought that the show was outrageous and left before the lights came on.	
1	fl	The woman that the hostile mechanic obviously patronized denied that the car was annoying and felt angry.	
1	fl	The intern that the psychology professor kindly dismissed had studied	

		under the terrible statistician for a while.	
1	f1	The teacher that the confused boy often questioned was teaching the Korean alphabet yesterday afternoon.	
1	f1	The captain that the smiling sailor warmly greeted was watching the beautiful sunset and steering the ship at the same time.	
1	f1	The customer that the computer technician actually helped forgot that the program was expensive.	
1	f1	The prince that the wicked sorceress quickly caught had saved the beautiful princess.	
1	f1	The housekeeper that the jealous butler loudly accused denied that the accountant was annoying.	
1	f1	The batter that the excited player constantly antagonized had swung at the tiny pitcher rudely.	
1	f2	The window reflected light so that her shadow was easily noticed by the man standing at the bookshelf.	Who noticed the shadow?
1	f2	The doctor treated the patients with compassion although there was little she could do to help.	Who treated the patient with compassion?
1	f2	The film director ordered the crew to be on the set at 6am but the leading man didn't show up until after lunch.	Who didn't show up?
1	f2	The board hired more general practitioners but the hospital decided they would replace the doctors with nurses instead.	Who hired more practitioners?
1	f2	The young reader anxiously ran towards the fiction section in the library yesterday morning.	Who ran towards the fiction section?
1	f2	The daring pilot dropped the seeds from the plane before he noticed the storm would come in a few minutes.	Who dropped the seeds?
1	f2	The student fell asleep at his desk during the class because he had rarely utilized the new computers.	Who fell asleep?
1	f2	The man had denied starting the illegal fire but he was seen hanging around the burnt building.	Who had denied starting the fire?
1	f2	The smart waitress complained to the manager that the salesman forgot that the drink was alcoholic.	Who complained?
1	f2	The couple had cleaned up after the rude game and they complained about this all night.	Who cleaned up after the game?
1	f2	The secretary answered the angry businessmen and was praised by the tired president last month.	Who answered the businessmen?
1	f2	The publicist had already paid for the brilliant painting but he wanted to cancel the exhibit now.	Who want to cancel the exhibit?
1	f2	The student and his concerned parents appreciated the help offered by	

		the advisor very much.	
1	f2	The mother dropped off the expensive bill and appreciated the discount very much just as the sales clerk expected.	
1	f2	The fan magazine reported that the actress was recognized at the market after she got the popular award.	
1	f2	The government admitted that the businesses were corrupted and refunded the taxpayer's money.	
1	f2	The boy disagreed that the movie was silly and argued with the sassy teenager angrily.	
1	f2	The traveler was convinced by the old gypsy to play the lottery someday.	
1	f2	The employer was threatened by the hostile worker because he accused the old employee.	
1	f2	The passenger was quickly followed by the FBI agent because he had hidden the dangerous explosive.	
1	f2	The biologist tested the contaminated water with the young zookeeper and left the zoo.	
1	f2	The queen offered the valiant prince a gorgeous castle located in a neighboring kingdom.	
1	f2	The boy was punished by his worried mother because he hit the little girl.	
1	f2	The property company hired someone less pushy to research the surrounding environment of the new building.	
2	f3	The dog with the bushy tail jumped all over him.	Who jumped?
2	f3	The woman with painted toenails did not wash her hand.	Who didn't wash her hands?
2	f3	The singer who was well-known thanked his fans during intermission.	Who thanked his fans?
2	f3	The large hospital with the budget problems fired the doctor.	
2	f3	The police inspector with the funny raincoat questioned the worker yesterday.	Who questioned the worker?
2	f3	The scenarios which are very complicated captured real circumstances.	
2	f3	The airport inspector who is very cautious interrogated the maid because of her rare passport.	Who interrogated the maid?
2	f3	The archeologist who came from Africa uncovered artifacts from a newly discovered site.	Who came from Africa?
2	f3	The student with the expensive bookbag looked like a spoiled brat.	
2	f3	The boy who was yelling ran to the top of the mountain.	
2	f3	Girls that wear purple tights and blue shirts make me laugh.	

2	f3	The email message that was deleted turned out to be a virus.	
2	f3	The popular bakery sold apple cakes which were very delicious.	Who sold apple cakes?
2	f3	The agent hoped that the customer who came from another city would take the fancy new house.	Who came from another city?
2	f3	The small country reduced the debt that caused the interest.	
2	f3	The first test included a variety of questions that demanded critical thinking.	
2	f3	The famous singer achieved the dream that she had been waiting for for so many years.	Who achieved the dream?
2	f3	The Supreme Court judge acquitted the man who claimed that the witness was lying.	Who was lying?
2	f3	The computer programmer adjusted the monitor that had been upgraded recently.	
2	f3	The lazy vacationer anticipated that the captain who discovered the island would explore the shore.	Who would explore the shore?
2	f3	The builders eliminated the idea that they should add another room onto the house.	Who eliminated the idea?
2	f3	The new scissors cut the wire that was keeping the door shut.	
2	f3	The large group played a game that kept them all entertained.	
2	f3	The fast runner set the pace that all his competitors kept up with.	
2	f4	The hungry fish swam quickly to the top of the tank.	
2	f4	The annoying alarm went off every nine minutes for half an hour.	
2	f4	The famous painter displayed her works at the local fine arts exhibition.	Who displayed the work?
2	f4	The informed citizen elected the most experienced candidate.	Who elected the candidate?
2	f4	The pool schedule said that the life guard would be on duty tonight.	Who was on duty tonight?
2	f4	The sports reporter interviewed the player before the championship game.	Who interviewed the player?
2	f4	My Aunt came home from England with many presents.	Who came back home?
2	f4	My sister was supposed to call me last night.	Who was supposed to call?
2	f4	The hospital received a lot of calls about the new cure.	
2	f4	The auctioneer sold the house because the owners moved to Florida.	Who sold the house?
2	f4	The shipping business kept records of all its clients that pay.	

2	f4	The woman sat alone because she was punished and had no regrets.	Who sat alone?
2	f4	The loving grandmother makes a cake for breakfast almost everyday.	Who makes cake everyday?
2	f4	The creative young architect preferred the complicated design.	Who preferred the complicated design?
2	f4	The ski-instructor warned the students of the icy conditions.	Who warned the students?
2	f4	The manager of the building was injured in a horrible car accident yesterday.	Who was injured?
2	f4	I ran ten miles today before I started swimming.	
2	f4	The principal of the high school suspended the undisciplined teens.	
2	f4	The cautious public doubted the commercials for the new product.	
2	f4	The state was suffering badly because of the unexpected hurricane.	
2	f4	The teacher was well respected among the board members.	
2	f4	The driver was arrested for drunk driving yesterday.	
2	f4	The financial consultant told the company to foresee every possible situation next year.	
2	f4	The father quickly threw the baseball to the child.	

Appendix B. Experimental Sentences

Item	cond	Sentence	Question
1	1	The client who had arrived after the important meeting that day was waiting in the office.	Who was waiting?
1	2	The client who had arrived after the important visitor that day was waiting in the office.	Who was waiting?
1	3	The client who implied that the meeting was important that day was waiting in the office.	Who was waiting?
1	4	The client who implied that the visitor was important that day was waiting in the office.	Who was waiting?
2	1	The resident who was living near the dangerous warehouse last month had complained about the investigation.	Who had complained?
2	2	The resident who was living near the dangerous neighbor last month had complained about the investigation.	Who had complained?
2	3	The resident who said that the warehouse was dangerous last month had complained about the investigation.	Who had complained?
2	4	The resident who said that the neighbor was dangerous last month had complained about the investigation.	Who had complained?
3	1	The teacher who was designing the new curriculum last night will come to the office.	Who will come?
3	2	The teacher who was meeting with the new specialist last night will come to the office.	Who will come?
3	3	The teacher who realized that the curriculum was new last night will come to the office.	Who will come?
3	4	The teacher who realized that the specialist was new last night will come to the office.	Who will come?
4	1	The ambassador who had exposed the known conspiracy during the meeting will arrive this morning.	Who will arrive?
4	2	The ambassador who had criticized the known terrorist during the meeting will arrive this morning.	Who will arrive?
4	3	The ambassador who claimed that the conspiracy was known in the meeting will arrive this morning.	Who will arrive?
4	4	The ambassador who claimed that the terrorist was known in the meeting will arrive this morning.	Who will arrive?
5	1	The critic who had enjoyed the memorable play at the new theater will praise the director.	Who will visit?
5	2	The critic who had enjoyed the memorable actress at the new theater	Who will visit?

		will praise the director.	
5	3	The critic who mentioned that the play was memorable at the new theater will praise the director.	Who will visit?
5	4	The critic who mentioned that the actress was memorable at the new theater will praise the director.	Who will visit?
6	1	The unit which was searching for the captured supplies after the bombing will contact the base.	Who will contact the base?
6	2	The unit which was searching for the captured pilots after the bombing will contact the base.	Who will contact the base?
6	3	The unit which suspected that the supplies were captured after the bombing will contact the base.	Who will contact the base?
6	4	The unit which suspected that the pilots were captured after the bombing will contact the base.	Who will contact the base?
7	1	The opponent who was fighting the corrupt government for nearly three year should be arrested immediately.	Who should be arrested?
7	2	The opponent who was fighting the corrupt governor for nearly three year should be arrested immediately.	Who should be arrested?
7	3	The opponent who had claimed that the government was corrupt for nearly three year should be arrested immediately.	Who should be arrested?
7	4	The opponent who had claimed that the governor was corrupt for nearly three year should be arrested immediately.	Who should be arrested?
8	1	The policeman who had found the missing money accidentally was expecting an investigation.	Who was expecting an investigation?
8	2	The policeman who had found the missing boy accidentally was expecting an investigation.	Who was expecting an investigation?
8	3	The policeman who had discovered that the money was missing accidentally was expecting an investigation.	Who was expecting an investigation?
8	4	The policeman who had discovered that the boy was missing accidentally was expecting an investigation.	Who was expecting an investigation?
9	1	The couple who had looked for the cheaper house since the wedding was making a mistake.	Who was making a mistake?
9	2	The couple who had looked for the cheaper agent since the wedding was making a mistake.	Who was making a mistake?
9	3	The couple who thought that the house was cheaper after the wedding was making a mistake.	Who was making a mistake?

9	4	The couple who thought that the agent was cheaper after the wedding was making a mistake.	Who was making a mistake?
10	1	The manager who liked the clever show at the opening ceremony could negotiate a good deal.	Who could negotiate?
10	2	The manager who liked the clever producer at the opening ceremony could negotiate a good deal.	Who could negotiate?
10	3	The manager who said that the show was clever at the opening ceremony could negotiate a good deal.	Who could negotiate?
10	4	The manager who said that the producer was clever at the opening ceremony could negotiate a good deal.	Who could negotiate?
11	1	The suspect who was aware of the unguarded money outside was sleeping during the crime.	Who was sleeping?
11	2	The suspect who was aware of the unguarded teller outside was sleeping during the crime.	Who was sleeping?
11	3	The suspect who knew that the money was unguarded outside was sleeping during the crime.	Who was sleeping?
11	4	The suspect who knew that the teller was unguarded outside was sleeping during the crime.	Who was sleeping?
12	1	The procedures which were favored by the designated laws in California were enacted throughout the country.	what were enacted?
12	2	The procedures which were favored by the designated voters in California were enacted throughout the country.	what were enacted?
12	3	The procedures which ensure that the laws are designated in California were enacted throughout the country.	what were enacted?
12	4	The procedures which ensure that the voters are designated in California were enacted throughout the country.	what were enacted?
13	1	The company which had created the controversial product last year asked for feedback.	Who asks for feedback?
13	2	The company which had hired the controversial spokesman last year asked for feedback.	Who asks for feedback?
13	3	The company that admited that the product is controversial last year asked for feedback.	Who asks for feedback?
13	4	The company that admited that the spokesman is controversial last year asked for feedback.	Who asks for feedback?
14	1	The student who was tried of the demanding assignment at school was leaving for home.	Who was leaving?
14	2	The student who was tried of the demanding coach at school was leaving for home.	Who was leaving?
14	3	The student who said that the assignment was demanding at school	Who was leaving?

		was leaving for home.	
14	4	The student who said that the coach was demanding at school was leaving for home.	Who was leaving?
15	1	The handyman who worked on the ambitious project for the company was standing outside.	Who was standing outside?
15	2	The handyman who worked with the ambitious foreman for the company was standing outside.	Who was standing outside?
15	3	The handyman who said that the project was ambitious to the company was standing outside.	Who was standing outside?
15	4	The handyman who said that the foreman was ambitious to the company was standing outside.	Who was standing outside?
16	1	The witness who will support the fraudulent case at the court will testify against the suspect.	Who will testify?
16	2	The witness who will support the fraudulent defendant at the court will testify against the suspect.	Who will testify?
16	3	The witness who suggested that the case was fraudulent at the court will testify against the suspect.	Who will testify?
16	4	The witness who suggested that the defendant was fraudulent at the court will testify against the suspect.	Who will testify?
17	1	The professor who agreed on the inappropriate reprimand yesterday was acting rudely.	Who was acting rudely?
17	2	The professor who agreed with the inappropriate chairman yesterday was acting rudely.	Who was acting rudely?
17	3	The professor who agreed that the reprimand was inappropriate yesterday was acting rudely.	Who was acting rudely?
17	4	The professor who agreed that the chairman was inappropriate yesterday was acting rudely.	Who was acting rudely?
18	1	The couple who had requested the sophisticated vase from the store was talking about the price.	Who was talking?
18	2	The couple who had requested a sophisticated salesman from the store was talking about the price.	Who was talking?
18	3	The couple who remarked that the vase was sophisticated at the store was talking about the price.	Who was talking?
18	4	The couple who remarked that the salesman was sophisticated at the store was talking about the price.	Who was talking?
19	1	The passenger who was sitting in the new seat on the bus was talking on the phone.	Who was talking?
19	2	The passenger who was sitting behind the new driver on the bus was talking on the phone.	Who was talking?

19	3	The passenger who commented that the seat was new on the bus was talking on the phone.	Who was talking?
19	4	The passenger who commented that the driver was new on the bus was talking on the phone.	Who was talking?
20	1	The physicist who had admired the amazing calculation at the conference was making too much noise.	Who was making noise?
20	2	The physicist who had admired the amazing chemist at the conference was making too much noise.	Who was making noise?
20	3	The physicist who shouted that the calculation was amazing at the conference was making too much noise.	Who was making noise?
20	4	The physicist who shouted that the chemist was amazing at the conference was making too much noise.	Who was making noise?
21	1	The attorney who was questioning the unusual motion in the courtroom was exaggerating quite a bit.	Who was exaggerating?
21	2	The attorney who was questioning the unusual witness in the courtroom was exaggerating quite a bit.	Who was exaggerating?
21	3	The attorney who commented that the motion was unusual in the courtroom was exaggerating quite a bit.	Who was exaggerating?
21	4	The attorney who commented that the witness was unusual in the courtroom was exaggerating quite a bit.	Who was exaggerating?
22	1	The candidate who was attacked by the dishonest commercial in the newspaper was losing the race.	Who was losing?
22	2	The candidate who was attacked by the dishonest senator in the newspaper was losing the race.	Who was losing?
22	3	The candidate who charged that the commercial was dishonest in the newspaper was losing the race.	Who was losing?
22	4	The candidate who charged that the senator was dishonest in the newspaper was losing the race.	Who was losing?
23	1	The publicist who had paid for the brilliant painting at the first meeting will cancel the exhibit.	Who will cancel the exhibit?
23	2	The publicist who had paid for the brilliant painter at the first meeting will cancel the exhibit.	Who will cancel the exhibit?
23	3	The publicist who assumed that the painting was brilliant at the first meeting will cancel the exhibit.	Who will cancel the exhibit?
23	4	The publicist who assumed that the painter was brilliant at the first meeting will cancel the exhibit.	Who will cancel the exhibit?
24	1	The judge who had criticized the questionable evidence recently had misunderstood the facts.	Who had misunderstood?
24	2	The judge who had criticized the questionable witness recently had	Who had

		misunderstood the facts.	misunderstood?
24	3	The judge who decided that the evidence was questionable recently had misunderstood the facts.	Who had misunderstood?
24	4	The judge who decided that the witness was questionable recently had misunderstood the facts.	Who had misunderstood?
25	1	The secretary who was complaining about the unreasonable policy on TV is quitting next month.	Who is quitting?
25	2	The secretary who was complaining about the unreasonable director on TV is quitting next month.	Who is quitting?
25	3	The secretary who complains that the policy is unreasonable on TV is quitting next month.	Who is quitting?
25	4	The secretary who complains that the director is unreasonable on TV is quitting next month.	Who is quitting?
26	1	The girl who was complaining about the annoying pain to everybody had visited the clinic.	Who had visited the clinic?
26	2	The girl who was complaining about the annoying mother to everybody had visited the clinic.	Who had visited the clinic?
26	3	The girl who said that the pain was annoying to everybody had visited the clinic.	Who had visited the clinic?
26	4	The girl who said that the mother was annoying to everybody had visited the clinic.	Who had visited the clinic?
27	1	The director who disliked the outrageous performance in the movie had wanted to quit.	Who wanted to quit?
27	2	The director who disliked the outrageous performer in the movie had wanted to quit.	Who wanted to quit?
27	3	The director who exclaimed that the performance was outrageous in the movie had wanted to quit.	Who wanted to quit?
27	4	The director who exclaimed that the performer was outrageous in the movie had wanted to quit.	Who wanted to quit?
28	1	The informant who had been exposing the illegal company to the public was arrested last night.	Who was arrested?
28	2	The informant who had been exposing the illegal immigrant to the public was arrested last night.	Who was arrested?
28	3	The informant who explained that the company was illegal to the public was arrested last night.	Who was arrested?
28	4	The informant who explained that the immigrant was illegal to the public was arrested last night.	Who was arrested?
29	1	The experts who were apologizing for the biased story during class will explain the decision.	Who will explain?

29	2	The experts who were apologizing for the biased judge during class will explain the decision.	Who will explain?
29	3	The experts who admitted that the story was biased during class will explain the decision.	Who will explain?
29	4	The experts who admitted that the judge was biased during class will explain the decision.	Who will explain?
30	1	The thief who had stolen from the strict church for a while lived near the sanctuary.	Who lived nearby?
30	2	The thief who had stolen from the strict nun for a while lived near the sanctuary.	Who lived nearby?
30	3	The thief who had known that the church was strict for a while lived near the sanctuary.	Who lived nearby?
30	4	The thief who had known that the nun was strict for a while lived near the sanctuary.	Who lived nearby?
31	1	The child who was playing with the dangerous toy all day long was running toward the park.	Who was running?
31	2	The child who was playing with the dangerous stranger all day long was running toward the park.	Who was running?
31	3	The child who wondered if the toy was dangerous all day long was running toward the park.	Who was running?
31	4	The child who wondered if the stranger was dangerous all day long was running toward the park.	Who was running?
32	1	The owner who had driven out the new undergrowth last summer had made mistakes.	Who had made mistakes?
32	2	The owner who had driven out the new assistant last summer had made mistakes.	Who had made mistakes?
32	3	The owner who regretted that the undergrowth was new last summer had made mistakes.	Who had made mistakes?
32	4	The owner who regretted that the assistant was new last summer had made mistakes.	Who had made mistakes?
33	1	The student who had merely seen the loud party last time was drinking under age.	Who was drinking?
33	2	The student who had merely seen the loud partygoer last time was drinking under age.	Who was drinking?
33	3	The student who testified that the party was loud last time was drinking under age.	Who was drinking?
33	4	The student who testified that the partygoer was loud last time was drinking under age.	Who was drinking?
34	1	The enemy who had shot at the hidden tank in the garage was lying	Who was lying in

		in the bushes.	the bushes?
34	2	The enemy who had shot at the hidden man in the garage was lying in the bushes.	Who was lying in the bushes?
34	3	The enemy who saw that the tank was hidden in the garage was lying in the bushes.	Who was lying in the bushes?
34	4	The enemy who saw that the man was hidden in the garage was lying in the bushes.	Who was lying in the bushes?
35	1	The burglar who had stolen the precious jewel from the store was scared by the alarm.	Who was scared?
35	2	The burglar who had stolen the precious baby from the store was scared by the alarm.	Who was scared?
35	3	The burglar who thought that the jewel was precious at the store was scared by the alarm.	Who was scared?
35	4	The burglar who thought that the baby was precious at the store was scared by the alarm.	Who was scared?
36	1	The salesman who had upset the alcoholic drink last night will complain to the manager.	Who will complain?
36	2	The salesman who had upset the alcoholic waiter last night will complain to the manager.	Who will complain?
36	3	The salesman who thought that the drink was alcoholic last night will complain to the manager.	Who will complain?
36	4	The salesman who thought that the waiter was alcoholic last night will complain to the manager.	Who will complain?
37	1	The doorman who had denounced the dreadful crime last month was planning something awful.	Who was planning something awful?
37	2	The doorman who had denounced the dreadful criminal last month was planning something awful.	Who was planning something awful?
37	3	The doorman who remarked that the crime was dreadful last month was planning something awful.	Who was planning something awful?
37	4	The doorman who remarked that the criminal was dreadful last month was planning something awful.	Who was planning something awful?
38	1	The hostess who had yelled about the dirty room loudly will forget about the mess.	Who will forget about the mess?
38	2	The hostess who had yelled about the dirty toddler loudly will forget about the mess.	Who will forget about the mess?
38	3	The hostess who yelled that the room was dirty loudly will forget about the mess.	Who will forget about the mess?
38	4	The hostess who yelled that the toddler was dirty loudly will forget about the mess.	Who will forget about the mess?

39	1	The secretary who had answered the angry calls this Monday should apologize to the company.	Who should apologize?
39	2	The secretary who had answered the angry businessmen this Monday should apologize to the company.	Who should apologize?
39	3	The secretary who claimed that the calls were angry this Monday should apologize to the company.	Who should apologize?
39	4	The secretary who claimed that the businessmen were angry this Monday should apologize to the company.	Who should apologize?
40	1	The teacher who had graded the difficult test last semester could answer the question.	Who could answer the question?
40	2	The teacher who had graded the difficult child last semester could answer the question.	Who could answer the question?
40	3	The teacher who disagreed that the test was difficult last semester could answer the question.	Who could answer the question?
40	4	The teacher who disagreed that the child was difficult last semester could answer the question.	Who could answer the question?
41	1	The teller who was working in the boring room in the back building will quit the job.	Who will quit?
41	2	The teller who was working for the boring boss in the back building will quit the job.	Who will quit?
41	3	The teller who felt that the room was boring in the back building will quit the job.	Who will quit?
41	4	The teller who felt that the boss was boring in the back building will quit the job.	Who will quit?
42	1	The subject who had argued about the difficult quiz on the school message board will leave soon.	Who will leave soon?
42	2	The subject who had argued about the difficult professor on the school message board will leave soon.	Who will leave soon?
42	3	The subject who learned that the quiz was difficult on the school message board will leave soon.	Who will leave soon?
42	4	The subject who learned that the professor was difficult on the school message board will leave soon.	Who will leave soon?
43	1	The waitress who had presented the greasy menu last time was unconcerned about the first impression.	Who was unconcerned?
43	2	The waitress who had presented the greasy cook last time was unconcerned about the first impression.	Who was unconcerned?
43	3	The waitress who knew that the menu was greasy last time was unconcerned about the first impression.	Who was unconcerned?
43	4	The waitress who knew that the cook was greasy last time was	Who was

		unconcerned about the first impression.	unconcerned?
44	1	The boy who had dumped the rich soil last week should apologize for the mistake.	Who should apologize?
44	2	The boy who had dumped the rich girl last week should apologize for the mistake.	Who should apologize?
44	3	The boy who ignored that the soil was rich last week should apologize for the mistake.	Who should apologize?
44	4	The boy who ignored that the girl was rich last week should apologize for the mistake.	Who should apologize?
45	1	The mother who had dropped off the expensive bill recently will appreciate the discount.	Who will appreciate the discount?
45	2	The mother who had dropped off the expensive teenager recently will appreciate the discount.	Who will appreciate the discount?
45	3	The mother who remarked that the bill was expensive recently will appreciate the discount.	Who will appreciate the discount?
45	4	The mother who remarked that the teenager was expensive recently will appreciate the discount.	Who will appreciate the discount?
46	1	The audience who was watching the outrageous show last night will remember the jokes.	Who will remember the jokes?
46	2	The audience who was watching the outrageous comedian last night will remember the jokes.	Who will remember the jokes?
46	3	The audience who thought that the show was outrageous last night will remember the jokes.	Who will remember the jokes?
46	4	The audience who thought that the comedian was outrageous last night will remember the jokes.	Who will remember the jokes?
47	1	The student who appreciated the helpful advise in the counseling session will make the right decision.	Who will make the right decision?
47	2	The student who appreciated the helpful advisor in the counseling session will make the right decision.	Who will make the right decision?
47	3	The student who mentioned that the advice was helpful in the counseling session will make the right decision.	Who will make the right decision?

47	4	The student who mentioned that the advisor was helpful in the counseling session will make the right decision.	Who will make the right decision?
48	1	The actress who had gotten the popular award in Europe was recognized in public.	Who was recognized?
48	2	The actress who had gotten the popular boyfriend in Europe was recognized in public.	Who was recognized?
48	3	The actress who loved that the award was popular in Europe was recognized in public.	Who was recognized?
48	4	The actress who loved that the boyfriend was popular in Europe was recognized in public.	Who was recognized?
49	1	The hostess who had ignored the messy table after the party was talking to some friends.	Who was talking?
49	2	The hostess who had ignored the messy guest after the party was talking to some friends.	Who was talking?
49	3	The hostess who ignored that the table was messy after the party was talking to some friends.	Who was talking?
49	4	The hostess who ignored that the guest was messy after the party was talking to some friends.	Who was talking?
50	1	The teacher who was teaching the Korean alphabet at school will translate the vocabulary.	Who will translate?
50	2	The teacher who was teaching the Korean child at school will translate the vocabulary.	Who will translate?
50	3	The teacher who knew that the alphabet was Korean at school will translate the vocabulary.	Who will translate?
50	4	The teacher who knew that the child was Korean at school will translate the vocabulary.	Who will translate?
51	1	The captain who was watching the beautiful sunset on the boat was steering the ship.	Who was steering?
51	2	The captain who was watching the beautiful woman on the boat was steering the ship.	Who was steering?
51	3	The captain who saw that the sunset was beautiful on the boat was steering the ship.	Who was steering?
51	4	The captain who saw that the woman was beautiful on the boat was steering the ship.	Who was steering?
52	1	The boy who was seeing the silly movie at the theater was wasting money.	Who was wasting money?
52	2	The boy who was seeing the silly girl at the theater was wasting money.	Who was wasting money?
52	3	The boy who disagreed that the movie was silly at the theater was	Who was wasting

		wasting money.	money?
52	4	The boy who disagreed that the girl was silly at the theater was wasting money.	Who was wasting money?
53	1	The customer who had asked about the expensive program yesterday should buy the computer.	Who should buy the computer?
53	2	The customer who had asked about the expensive programmer yesterday should buy the computer.	Who should buy the computer?
53	3	The customer who forgot that the program was expensive yesterday should buy the computer.	Who should buy the computer?
53	4	The customer who forgot that the programmer was expensive yesterday should buy the computer.	Who should buy the computer?
54	1	The prince who had saved the beautiful castle last time had forgotten the magic wand.	Who had forgotten the wand?
54	2	The prince who had saved the beautiful princess last time had forgotten the magic wand.	Who had forgotten the wand?
54	3	The prince who thought that the castle was beautiful last time had forgotten the magic wand.	Who had forgotten the wand?
54	4	The prince who thought that the princess was beautiful last time had forgotten the magic wand.	Who had forgotten the wand?
55	1	The housekeeper who had told about the secret room to the media was fired for the error.	Who was fired?
55	2	The housekeeper who had told about the secret accountant to the media was fired for the error.	Who was fired?
55	3	The housekeeper who revealed that the room was a secret to the media was fired for the error.	Who was fired?
55	4	The housekeeper who revealed that the accountant was a secret to the media was fired for the error.	Who was fired?
56	1	The batter who had swung at the tiny ball at the last game should play baseball.	Who should play?
56	2	The batter who had swung at the tiny pitcher at the last game should play baseball.	Who should play?
56	3	The batter who said that the ball was tiny at the last game should play baseball.	Who should play?
56	4	The batter who said that the pitcher was tiny at the last game should play baseball.	Who should play?

57	1	The traveler who had dreamt about the dark shadow last night will win the lottery.	Who will win?
57	2	The traveler who had dreamt about the dark man last night will win the lottery.	Who will win?
57	3	The traveler who remembered that the shadow was dark last night will win the lottery.	Who will win?
57	4	The traveler who remembered that the man was dark last night will win the lottery.	Who will win?
58	1	The employer who had criticized the old computer at the company meeting will regret the nasty comment.	Who will regret?
58	2	The employer who had criticized the old employee at the company meeting will regret the nasty comment.	Who will regret?
58	3	The employer who remarked that the computer was old at the company meeting will regret the nasty comment.	Who will regret?
58	4	The employer who remarked that the employee was old at the company meeting will regret the nasty comment.	Who will regret?
59	1	The family which had enjoyed the offensive music at the concert was ignorant about quality performances.	Who was ignorant?
59	2	The family which had enjoyed the offensive singer at the concert was ignorant about quality performances.	Who was ignorant?
59	3	The family which had disregarded that the music was offensive at the concert was ignorant about quality performances.	Who was ignorant?
59	4	The family which had disregarded that the singer was offensive at the concert was ignorant about quality performances.	Who was ignorant?
60	1	The passenger who had hidden the dangerous explosive in the back room will leave immediately.	Who will leave immediately?
60	2	The passenger who had hidden the dangerous criminal in the back room will leave immediately.	Who will leave immediately?
60	3	The passenger who understood that the explosive was dangerous in the back room will leave immediately.	Who will leave immediately?
60	4	The passenger who understood that the criminal was dangerous in the back room will leave immediately.	Who will leave immediately?
61	1	The biologist who was testing the contaminated water in the village was afraid of dying.	Who was afraid?
61	2	The biologist who was testing the contaminated animal in the village was afraid of dying.	Who was afraid?
61	3	The biologist who disliked that the water was contaminated in the village was afraid of dying.	Who was afraid?
61	4	The biologist who disliked that the animal was contaminated in the	Who was afraid?

		village was afraid of dying.	
62	1	The mother who had punished the bad dog last week will explain the reason.	Who will explain the reason?
62	2	The mother who had punished the bad boy last week will explain the reason.	Who will explain the reason?
62	3	The mother who thought that the dog was bad last week will explain the reason.	Who will explain the reason?
62	4	The mother who thought that the boy was bad last week will explain the reason.	Who will explain the reason?
63	1	The hero who had saved the frightened cat from the burning car could explain the accident.	Who could explain the accident?
63	2	The hero who had saved the frightened lady from the burning car could explain the accident.	Who could explain the accident?
63	3	The hero who saw that the cat was frightened of the burning car could explain the accident.	Who could explain the accident?
63	4	The hero who saw that the lady was frightened of the burning car could explain the accident.	Who could explain the accident?
64	1	The boy who had stolen the old trophy that night should stop stealing.	Who should stop stealing?
64	2	The boy who had stolen from the old man that night should stop stealing.	Who should stop stealing?
64	3	The boy who knew that the trophy was old that night should stop stealing.	Who should stop stealing?
64	4	The boy who knew that the man was old that night should stop stealing.	Who should stop stealing?
65	1	The captain who had discovered the incredible island eventually will explore more.	Who will explore more?
65	2	The captain who had discovered the incredible woman eventually will explore more.	Who will explore more?
65	3	The captain who realized that the island was incredible eventually will explore more.	Who will explore more?
65	4	The captain who realized that the woman was incredible eventually will explore more.	Who will explore more?
66	1	The manager who had lied about the cheap car in the end will resign before the summer.	Who will resign?

66	2	The manager who had lied about the cheap associate in the end will resign before the summer.	Who will resign?
66	3	The manager who admitted that the car was cheap in the end will resign before the summer.	Who will resign?
66	4	The manager who admitted that the associate was cheap in the end will resign before the summer.	Who will resign?
67	1	The bandit who had attacked the frail wagon last night was looking for help.	Who was looking for help?
67	2	The bandit who had attacked the frail woman last night was looking for help.	Who was looking for help?
67	3	The bandit who noticed that the wagon was frail last night was looking for help.	Who was looking for help?
67	4	The bandit who noticed that the woman was frail last night was looking for help.	Who was looking for help?
68	1	The worker who had picked up the lost ticket from the front desk will come to the doctor.	Who will come to the doctor?
68	2	The worker who had picked up the lost child from the front desk will come to the doctor.	Who will come to the doctor?
68	3	The worker who saw that the ticket was lost at the front desk will come to the doctor.	Who will come to the doctor?
68	4	The worker who saw that the child was lost at the front desk will come to the doctor.	Who will come to the doctor?
69	1	The businessman who was complaining about the rude delay this afternoon should call the airline.	Who should call?
69	2	The businessman who was complaining about the rude passenger this afternoon should call the airline.	Who should call?
69	3	The businessman who complained that the delay was rude this afternoon should call the airline.	Who should call?
69	4	The businessman who complained that the passenger was rude this afternoon should call the airline.	Who should call?
70	1	The merchant who had sold the creative artwork in the exhibition was lying about the price.	Who was lying?
70	2	The merchant who had sold to the creative artist in the exhibition was lying about the price.	Who was lying?
70	3	The merchant who said that the artwork was creative in the exhibition was lying about the price.	Who was lying?
70	4	The merchant who said that the artist was creative in the exhibition was lying about the price.	Who was lying?
71	1	The player who was practicing on the new field before the game will	Who will score

		score many points.	many points?
71	2	The player who was practicing with the new team before the game will score many points.	Who will score many points?
71	3	The player who realized that the field was new before the game will score many points.	Who will score many points?
71	4	The player who realized that the team was new before the game will score many points.	Who will score many points?
72	1	The conservationist who had supported the pleasant park this year was speaking at the forum.	Who was speaking?
72	2	The conservationist who had supported the pleasant politician this year was speaking at the forum.	Who was speaking?
72	3	The conservationist who remarked that the park was pleasant this year was speaking at the forum.	Who was speaking?
72	4	The conservationist who remarked that the politician was pleasant this year was speaking at the forum.	Who was speaking?
73	1	The editor who had interviewed for the prestigious position yesterday was receiving a reward.	Who was rewarded?
73	2	The editor who had interviewed the prestigious citizen yesterday was receiving a reward.	Who was rewarded?
73	3	The editor who recognized that the position was prestigious yesterday was receiving a reward.	Who was rewarded?
73	4	The editor who recognized that the citizen was prestigious yesterday was receiving a reward.	Who was rewarded?
74	1	The child who was inspired by the creative assignment this summer was thankful for the attention.	Who was thankful?
74	2	The child who was inspired by the creative teacher this summer was thankful for the attention.	Who was thankful?
74	3	The child who said that the assignment was creative this summer was thankful for the attention.	Who was thankful?
74	4	The child who said that the teacher was creative this summer was thankful for the attention.	Who was thankful?
75	1	The boy who had played the interesting game last week had shared the cookies.	Who had shared the cookies?
75	2	The boy who had played with the interesting visitor last week had shared the cookies.	Who had shared the cookies?
75	3	The boy who mumbled that the game was interesting last week had shared the cookies.	Who had shared the cookies?
75	4	The boy who mumbled that the visitor was interesting last week had shared the cookies.	Who had shared the cookies?

76	1	The specialist who has worked with the terminal disease in the hospital will discuss all the options.	Who will discuss all the options?
76	2	The specialist who has worked with the terminal patient in the hospital will discuss all the options.	Who will discuss all the options?
76	3	The specialist who had thinks that the disease is terminal in the hospital will discuss all the options.	Who will discuss all the options?
76	4	The specialist who had thinks that the patient is terminal in the hospital will discuss all the options.	Who will discuss all the options?
77	1	The monster who had captured the precious sword in the battle will regret the action.	Who will regret the action?
77	2	The monster who had captured the precious queen in the battle will regret the action.	Who will regret the action?
77	3	The monster who realized that the sword was precious in the battle will regret the action.	Who will regret the action?
77	4	The monster who realized that the queen was precious in the battle will regret the action.	Who will regret the action?
78	1	The mailman who was avoiding the angry dog last time will knock on the door.	Who will knock?
78	2	The mailman who was avoiding the angry woman last time will knock on the door.	Who will knock?
78	3	The mailman who knew that the dog was angry last time will knock on the door.	Who will knock?
78	4	The mailman who knew that the woman was angry last time will knock on the door.	Who will knock?
79	1	The critic who had liked the interesting painting for years will buy something.	Who will buy something?
79	2	The critic who had liked the interesting painter for years will buy something.	Who will buy something?
79	3	The critic who had thought that the painting was interesting for years will buy something.	Who will buy something?
79	4	The critic who had thought that the painter was interesting for years will buy something.	Who will buy something?
80	1	The cashier who had screamed about the dangerous fire in the lobby was looking for the exit	Who was looking for the exit?
80	2	The cashier who had screamed about the dangerous robber in the lobby was looking for the exit	Who was looking for the exit?
80	3	The cashier who screamed that the fire was dangerous in the lobby was looking for the exit	Who was looking for the exit?
80	4	The cashier who screamed that the robber was dangerous in the lobby	Who was looking

		was looking for the exit	for the exit?
		Fillers	
f1	51	The girl who the seamstress reported on the plane had searched for the unique fabric over and over again.	Who reported that girl?
f1	51	The boy who the neighbor found outside was sent back home. immediately	Who found the boy?
f1	51	blank The queen who the jester amused easily realized that the singer was popular.	Who was popular?
f1	51	The student who the board acquitted officially had visited the loud party that night.	Where had the student visited?
f1	51	blank The player who the coach supported after the accident had played for many years.	Who supported the player
f1	51	blank The surgeon who the doctor knew in the new hospital has operated on many patients.	Where did the doctor know the surgeon?
f1	51	blank The assassin who the bodyguard noticed in the crowd had shot at the governor.	Who had been shot?
f1	51	The physicist who the scientist admired in the promising scientific area made too much noise in his office.	Where did the physicist make noise?
f1	51	The teacher who the girl trusted all the times had unfairly graded the test this time.	How did the teacher grade the test?
f1	51	blank The company who the customer criticized constantly figured out that the technician was unreliable.	Who was unreliable?
f1	51	blank The waitress who the driver saw outside of the restaurant knew that the menu was greasy.	Who saw the waitress?
f1	51	The boy who the counselor advised at the end of the semester regretted his ignorance on the subject.	What did the boy regret?
f1	51	The customer who the grocer checked out just now cried that the bag was broken loudly.	What was heavy?
f1	51	The child who the teacher ignored temperately complained that the tape was loud repeatedly.	How often did the child complain?
f1	51	The mother who the daughter reproached in private paid for the filthy dress happily.	How did the mother pay for the dress?
f1	51	blank The friend who the brothers admired at school left before the lights came on.	Where did the brother criticize the man?

f2	52	The window reflected light so that the lady's shadow was easily noticed by the visitor.	Whose shadow was easily notice?
f2	52	The doctor treated the patients although there was little she could do to help with the symptom.	Who treated the patients?
f2	52	The director ordered the crew to be on the set but the actor didn't show up this morning.	Who didn't show up?
f2	52	The board hired more practitioners recently but the hospital decided to replace the doctors.	Who did the board hire recently?
f2	52	blank The reader anxiously ran towards the fiction section in the library and started reading.	Who started reading?
f2	52	The pilot dropped the seeds before he noticed that the storm would come at any moment.	What would come at any moment?
f2	52	blank The student fell asleep during the class because he had stayed up last night.	When did the student stay up?
f2	52	The defendant had denied starting the fire but he was seen hanging around at that time.	Who had denied?
f2	52	blank The waitress complained to the manager that he forgot that the drink was alcoholic.	Who did the waitress complain to?
f2	52	blank The couple cleaned up the table after the game and left late.	What did the couple clean up?
f2	52	blank The secretary solved the crisis and was praised by the president in public.	Who praised the secretary?
f2	52	The investor had already paid for the painting but he wanted the refund afterward.	What did the investor want afterward?
f2	52	The student and his parents appreciated the help offered by the advisor.	Who offered the help?
f2	52	The mother dropped off the expensive bill and decided to switch to another service provider.	What did the mother dropped off?
f2	52	The newspaper reported that the actor was chosen by the director based on his performance in the last movie.	Who was chosen?
f2	52	The government admitted that the businesses were corrupted and promised to improve the situation as soon as possible.	What would be improved?
f3	53	The dog with the bushy tail jumped all over the trainer after the door got open.	Who did the dog jump all over?
f3	53	blank The popular bakery sold the popular apple cake which was very delicious.	Who sold the cake?

f3	53	blank The computer programmer adjusted the monitor that had been upgraded recently.	When had the monitor been upgraded?
f3	53	The judge acquitted the man who claimed that the witness was lying at the court.	Who acquitted the man
f3	53	The company which knew that the manager was annoying would fire the employee.	Who was annoying?
f3	53	The builders eliminated the idea that they should add another room to the new house.	What did the builders eliminate?
f3	53	blank The email message that was deleted turned out to be a virus after all.	What did the email turned out to be?
f3	53	blank The small country reduced the debt that caused the interest.	What caused the interest?
f3	53	blank The first test included a variety of questions that demanded critical thinking.	What did the questions demand?
f3	53	blank The thrifty husband left the woman who the owner liked.	Who left the woman?
f3	53	blank The scenarios which were very complicated captured the real circumstance.	What did the scenario capture?
f3	53	The skater really hated the manager who the coach rejected hard.	Who rejected the manager?
f3	53	blank The police inspector with the funny raincoat questioned the worker yesterday.	When was the worker questioned?
f3	53	blank The magician that tricked the lady at the show was the most popular one at that time.	Who was tricked?
f3	53	The new scissors cut the wire that was keeping the door shut all the time.	What was cut?
f3	53	The advice that was helpful for the student prevented the careless behavior in his life.	Who was the advice helpful for?
f4	54	blank The hungry fish swam quickly to the top of the tank.	How fast did the fish swim?
f4	54	blank The annoying alarm went off every nine minutes for half an hour everyday.	What went off?
f4	54	blank The famous painter displayed her works last weekend at the	When did the

		local fine arts exhibition.	painter display her works?
f4	54	blank The informed citizen elected the most experienced candidate in the recent election.	Who elected the candidate
f4	54	blank The pool manager said that the life guard would be on duty tonight.	When would the life guard be on duty?
f4	54	blank The sport reporter interviewed the player briefly before the championship game.	How did the reporter interview the player?
f4	54	blank The hospital received a lot of calls about the new cure published recently.	What were the calls about?
f4	54	The auctioneer sold the house because the owners moved to Florida.	Who moved to Florida?
f4	54	The kid sat alone in the lobby because she was punished and had no regrets.	Where did the woman sit?
f4	54	blank The loving grandmother makes a cake for breakfast almost everyday.	What does the grandmother make?
f4	54	blank The creative young architect preferred the complicated design of the new theater after the serious consideration.	What kind of design did the architect prefer?
f4	54	The ski-instructor warned the students of the icy conditions in the beginning of the training.	Who warned the students?
f5	55	The pilot who admired the captain greatly will argue with the stewardess.	Who would the pilot argue with?
f5	55	The police officer who complained to the man about the incident was expecting a reward.	What did the police officer complain about?
f5	55	The doctor who prescribed the horrible medicine for the patient had seen the results.	What was prescribed for the patient?
f5	55	The fashion editor who worked for the company on this promising project could discontinue the unusual color.	What could be discontinued?
f5	55	The neighbor who waked up the old couple accidentally was contacted by the residential committee.	Who contacted the neighbor?
f5	55	The salesman who tried to sell the house to the bachelor made a mistake.	What did the salesman try to sell?

f5	55	The seamstress who misled the girl that day told her a lie.	Who was told to be exclusive?
f5	55	The queen who asked for the popular song at the concert enjoyed the show.	What did the queen enjoy?
f5	55	The coach who supported that the competition was tough all the time listened for directions.	What was tough?
f5	55	blank The consultant who observed that the students had rarely utilized the new computer was surprised.	What had the students rarely utilized?
f5	55	The employee who investigated the agency undercover found them unreliable.	Who were unreliable?
f5	55	The artist who feared that the publicist would cancel the exhibit in the end quitted on his own.	What might be cancelled?
f5	55	Girls who wear purple tights and blue shirts in the show make me laugh.	Who laughs?
f5	55	The boy who yelled that the puppy was getting away loudly ran to the top of the mountain.	Which part of the mountain did the boy run to?
f5	55	The customer who said that his parents came from another city would take the fancy house.	What would the customer take?
f5	55	The inspector who noticed that the passport was rare in the investigation interrogated the maid.	What was rare?
f5	55	The archeologist who learned that an ancient trade team passed here six hundred years ago uncovered the precious artifacts.	Who passed here six hundred years ago?
f5	55	The singer who knew that the song was popular in the seventies thanked his fans.	What was popular?
f5	55	blank My sister who supposed that I would call her after school got disappointed.	Who was supposed to call?
f5	55	The father who just came back from work quickly threw the baseball to the child.	Who was the baseball thrown to?

Appendix C. Full correlation matrices of the correlation tests

1) Full correlation matrix of the correlation tests between interference effect size and individual differences measures.

Effect	Regions		Operation	Reading	Vocabulary	Category	Stroop	Digit	SAT
Syn	Accuracy	*r*	-0.002	-0.057	0.035	0	0	0.082	0.135
		p	0.987	0.582	0.735	0.999	0.993	0.43	0.191
Sem	Accuracy	*r*	-.207*	-.243*	0.104	-0.03	0.026	-0.008	-0.036
		p	0.043	0.017	0.312	0.773	0.804	0.941	0.73
RT									
Syn	Critical	*r*	0.045	-0.026	0.058	0.025	-0.091	0.036	0.015
		p	0.662	0.804	0.575	0.811	0.376	0.724	0.883
Sem	Critical	*r*	-0.02	-0.13	-0.042	-.239*	-0.11	-0.029	0.068
		p	0.849	0.206	0.686	0.02	0.287	0.782	0.509
Syn	Spillover	*r*	-0.093	-.231*	0.005	-0.092	0.05	-0.018	-.221*
		p	0.366	0.023	0.96	0.376	0.629	0.86	0.03
Sem	Spillover	*r*	0.028	-0.083	-0.13	-0.073	0.115	-0.001	-0.094
		p	0.79	0.425	0.21	0.486	0.268	0.99	0.363
Syn	Question	*r*	0.004	0.047	0.086	0.073	-0.159	-0.009	0.125
		p	0.969	0.649	0.408	0.482	0.123	0.934	0.226
Sem	Question	*r*	0.005	-0.177	-.260*	-.267**	.204*	-0.093	-.261*
		p	0.964	0.088	0.011	0.01	0.048	0.374	0.011

Note. "Syn" represent "syntactic interference effect size" and "Sem" represents "semantic interference effect size".

2) Full correlation matrix of the correlation tests between individual differences measures.

		Operation	Reading	Vocabulary	Category	Stroop	Digit	SAT
Operation	*r*	1	.510**	.229*	.224*	-.254*	.215*	0.092
	p		0	0.025	0.029	0.012	0.036	0.371
Reading	*r*	.510**	1	.361**	.326**	-0.161	0.174	.347**
	p	0		0	0.001	0.116	0.091	0.001
Vocabulary	*r*	.229*	.361**	1	.222*	-0.171	0.179	.606**
	p	0.025	0		0.031	0.095	0.081	0
Category	*r*	.224*	.326**	.222*	1	-0.031	0.147	.222*
	p	0.029	0.001	0.031		0.769	0.156	0.031
Stroop	*r*	-.254*	-0.161	-0.171	-0.031	1	-0.091	-.246*
	p	0.012	0.116	0.095	0.769		0.38	0.016
Digit	*r*	.215*	0.174	0.179	0.147	-0.091	1	.211*
	p	0.036	0.091	0.081	0.156	0.38		0.039
SAT	*r*	0.092	.347**	.606**	.222*	-.246*	.211*	1
	p	0.371	0.001	0	0.031	0.016	0.039	